#1-3

LEHIGH UNIVERSITY STUDIES

Psychology Series, No. 1 — 3

Learning to Generalize
and
An Experiment in Generalizing: A Unicursal Problem

by

JAMES L. GRAHAM, Ph.D.
Associate Professor of Psychology

Reprinted from Psychological Monographs and the Journal of Experimental Psychology

LEHIGH UNIVERSITY
BETHLEHEM, PENNSYLVANIA
1939

LEHIGH UNIVERSITY PUBLICATION

Published monthly during the calendar year by Lehigh University, Bethlehem, Pennsylvania. Entered as second-class matter, March 24, 1927, at the Post Office at Bethlehem, Pennsylvania, under the Act of August 24, 1912.

Circular No.	Title
*1.	Organization of the Institute of Research of Lehigh University.
*2.	Microscopical Studies of Anthracite.
3.	Rate of Molecular Weight Increase in the Boiling of Linseed Oil.
*4.	A Preliminary Study of Magnesium-Base Alloys.
*5.	Absorption of Carbon Dioxide by Coal.
6.	Studies in the Boiling of Linseed and China Wood Oils. 25 cents.
7.	The Shakespeare Folios and the Forgeries of Shakespeare's Handwriting in the Lucy Packer Linderman Memorial Library at Lehigh University. 25 cents.
*8.	Mazzini and Dante. 20 cents.
9.	The Center, Function and Structure of Psychology. 15 cents
10.	Methods for Exciting and for Calibrating Tuning Forks. 15 cents.
*11.	Flat Luminous Flames. 10 cents.
12.	The Pressure Vaccination Technic. 10 cents.
13.	Studies in Drying Oils. 10 cents.
14.	The Variant Issues of Shakespeare's Second Folio and Milton's First Published English Poems. 25 cents.
*15.	The Hydrates of Calcium Nitrate. 10 cents.
*16.	Tower Absorption Coefficients. 10 cents.
*17.	The Testing of Audio-Frequency Transformer-Coupled Amplifiers. 25 cents.
*18.	Neurovaccine. 10 cents.
19.	The Action of Bromine on Unsaturated Fatty Acids. 10 cents
*20.	Equal-Slope Surfaces and Helices by Vectors. 10 cents.
*21.	Coal Conductivity Cell. 10 cents.
*22.	Volume Changes During Hydration of Gelatin, Cement and Plaster of Paris. 10 cents.
23.	Studies in Drying Oils. 10 cents.
24.	Mononitro- and Dinitrothiophenes. 10 cents.
25.	Studies in Flame Propagation. 10 cents.
26.	The Biochemistry of the Soaking of Hides. (Parts 1 and 2.) 20 cents.
*27.	Petroleum Wash-Oil Thickening in the Scrubbing of Coke-Oven Gas. 10 cents.
28.	Government. A Phase of Social Organization. 50 cents.
29.	Aspects of Slavery and Expansion. 50 cents.
30.	Lipid Distribution in Normal and Abnormal Liver Tissues. 10 cents.
31.	Studies in Drying Oils. 10 cents.
32.	Scientific Papers from Department of Physics. 10 cents.
33.	Scientific Papers from the Department of Mathematics. 25 cents.
*34.	Effect of Nitrate Oxygen upon Tannery Effluent. 10 cents.
35.	Contsitution and Nature of Pennsylvania Anthracite. 10 cents.
36.	Scientific Papers from Department of Physics, 1929. 10 cents
37.	Studies in Drying Oils. 10 cents.
*38.	The Biochemistry of Soaking of Hides. 20 cents.
39.	Papers from the Department of Metallurgy. 25 cents.
40.	Scientific Papers from Department of Mathematics for 1929. 25 cents.
41.	Floor Test in the George Mason Hotel, Alexandria, Va., 15 cents.
*42.	Investigation of Welded Connections between Beams and Columns. 15 cents.
*43.	Forms of Generalization and Their Causes. 10 cents.
*44.	Thermal Treatment of Natural Gas. 10 cents.
*45.	A Proof that the Induction Motor Circle Diagram Applies to the Transmission Line. 10 cents.
*46.	Studies in Drying Oils. 10 cents.
47.	The Biochemistry of the Soaking and Liming of Hides. 10 cents.
48.	The Hydration of Animal Skin by the Volume Change Method. 10 cents.
*49.	Studies of Some Properties of Gelatin. 10 cents.
*50.	Rate of Oxidation of Linseed Oil at 160° C. 10 cents.
51.	Shakespeare Allusions and Parallels. 50 cents.
52.	Compressive Strength of Concrete in Flexure as Determined from Tests of Reinforced Beams. 20 cents.
*53.	Energy of the Iron Arc. 10 cents.
54.	The Relative Merits of Some Different Alloy Steels with Respect to Certain Mechanical Properties. 10 cents.
*55.	The United States Indian Policy in Texas. 10 cents.
*56.	Logic as the Cross-Classification and Selection of Arbitrary Elements. 10 cents.
57.	Studies in Drying Oils. 10 cents
58.	Heats of Wetting and Adsorption on Zinc Oxide. 10 cents.
*59	Some Experiments on the Soaking of Silks. 10 cents.
*60.	The Stabilization of Blue Cupric Hydroxide. 10 cents.
*61.	A Study of Tannery Effluent. 10 cents.
*62.	Hydration of Animal Skin by Volume Change Method. 10 cents
*63.	A Study of the Liming and Depilation of Animal Skin. 10 cents.
64.	Tests of Reinforced Concrete Columns. 50 cents.
*65.	Studies in Colonial Connecticut Taxation. 20 cents
*66.	The Age of Roman Sacrificial Victims. 10 cents.
*67.	X-Ray Analysis of Slate. 10 cents.
*68	A Critical Study of Pickling. 20 cents.
*69.	Gamma Ray Testing of Welds. 10 cents.
70.	Silver—A Senate Racket. 10 cents.
*71.	Scientific Papers from the Department of Mathematics. 20 cents.
*72.	X-Ray Studies of Animal Skin. I. Structure of Calf Skin as Indicated by Molybdenum Radiation. 10 cents.

(List continued on inside back cover)

[Reprinted from the Peterson Memorial Number, PSYCHOLOGICAL MONOGRAPHS, 1938, Vol. 50, Peabody Studies in Psychology.]

LEARNING TO GENERALIZE

JAMES L. GRAHAM
LEHIGH UNIVERSITY

In generalization problems, the learner's operations are essentially mental; the preliminaries to relational discriminations are usually submerged in chaotic frustrations; and the progress in the earlier stages of the learning can not be directly scored. These difficulties militate against explaining the learning process and skills involved.

This study attempts to provide exercises for generalizing, sufficiently abbreviated to come within reasonable time and motivational limits, and to examine, first, the nature of the tasks, second, some of the skill components involved in successful generalizing, and third, the abilities of college students and the relation of these abilities to certain products of learning. The interest is in presenting a theoretical analysis of this very broad field, based upon the assumption that the generalizing task should be considered as a form of learning. (For a contrasting position, see Maier [23]). The magnitude of this task and the subjective nature of the processes will necessitate the omission of historical material and experimental detail.

THE NATURE OF THE TASKS

The tasks could be completely solved by the learning of one or of a few simple relationships which reveal how to get answers for a whole series of problems. They are to be differentiated in whole or in part from related tasks of the following types: those which require the solution of problems by assigning specific answers (24, 33), those involving combination or reorganization of elements or experiences (15, 20, 22), the deduction of useful formulations by the manipulation of accepted statements in accordance with known rules (26), the abstraction of a common uniformity from varying specific elements or experiences (10, 11),

the delayed recall or recognition of previously acquired generalizations or categories of knowledge (38), the application of a known relationship to some new situation (3, 35, 38), and the shift from an accustomed or suggested inhibiting set to a favorable one (19, 21). The general pattern of the problems, materials, directions, sets, and suggestions given as instructions, are described as follows:

1. The instructions required the learning of a method of obtaining answers, from the study of certain specific situations and data. Usually, a formula or verbal description of the derived method of procedure satisfied this goal of learning.

2. An attempt was made to provide both materials for manipulation and an adequate description of the whole problem situation. Typically, this involved a description of a game, materials used, governing rules, methods of play, objectives, and scoring possibilities or comparable equivalents.

3. In order to reduce the number of steps to be learned, and to prevent observational errors which would inhibit the discernment and checking of relationships, specific answers within an arbitrarily selected range, or comparable information regarding the problems were supplied.

4. Specific answers were usually grouped in a more or less systematic arrangement and, typically, involved the presentation of answers in a continuous series beginning with simple forms.

5. In some of the tasks, further arranging, grouping, and other manipulating procedures were suggested, to give the learner an opportunity to observe relationships which were not apparent, or to present them in a different context.

A detailed description of two of these generalizing problems is included in the Appendix to this article (Section A), and a brief description or reference is made to additional generalizations or situations. Some of these are familiar to psychologists in other forms of presentation. The nature of the modification can be inferred from the detailed examples given and the general pattern outlined above. The familiar water-dipping "Ingenuity Test" of the Stanford-Binet Intelligence Scale (36, 37) can be taken for further illustration. The materials furnished are sand

and a graded series of vessels. The types are separated, a graduated series of answers is constructed, and appropriate generalized procedures are required for each. The generalization for the simplest type may be described as manipulating to retain the difference between the large and small vessels. A sample answer series may begin: To measure exactly 3 units, using vessels 4 and 7; fill 4 from 7, and retain 3 in 7.

An accurate understanding of differences in problem-solving tasks is aided by a careful description and analysis of the special conditions imposed by the type of response required, by the kinds and amounts of change necessitated, by the instructions, and by the apparatus used (30, p. 387 ff.). The processes involved in conserving, retaining, or applying *old* knowledge, and those involved in discriminating and acquiring *new* or revising the old, may differ enough to justify the development of specialized meanings or new terms, as Maier has attempted to do by restricting learning to situations more like the former, and substituting reasoning for problem solving situations more like the latter (23). In any case, the special conditions involved need elucidation, for both the situation and the response aspects.

I turn, then, to five general aspects of learning, to examine them in relation to certain special conditions imposed by my tasks. These aspects are: the learner's starting level, the objective or goal of the learning, the possibilities or limitations of scoring the progress of the learning, the amount of modification necessary to meet the criterion of the learning, and the probable transfer to similar tasks of the sets and skills used.

1. *Special Conditions Related to the Starting Level.* This starting level includes all that the learner brings to the start of the learning, such as: the levels of maturation, intelligence, and familiar knowledge, as well as the skills, habits, attitudes and earlier experience of job satisfaction or dissatisfaction. With reference to the suitability of my tasks and their bearing on where the learner starts, two questions are discussed: (A) What is the relation between the college student's large stock of acquired generalizations and the appropriateness of my tasks to his abili-

ties? (B) What is the relation of his accomplishment to his habituated sets and techniques?

(A) It is difficult to distinguish between new acquisition and recall of previous learning (to tell, for example, whether or not some of the generalizations in the water-dipping problem require new acquisitions or merely the recall of known principles). The criteria I used for determining unfamiliarity included a lack of school instruction upon the generalizations or situations, the student's reports of unfamiliarity, and even a low percentage of successful performance. The finding of suitable but unfamiliar generalizing tasks for college students presents many difficulties, since most relationships which are found in simple settings and which have practical value have been learned. The most suitable problems are low in utilitarian value, and therefore, provide little intrinsic motivation. Furthermore, such tasks are difficult (or give the initial impression of extrême complexity), and hence are unduly disheartening to the learner. Finally tasks for which a complete solution may be derived, in the sense that one or a few relationships will describe how to obtain all correct answers and no others, are mathematical in type, and many individuals have previously developed unfavorable attitudes toward manipulating such symbols and concepts.

(B) The importance of the rôle of habituated sets and attitudinal predispositions is evident. Two examples taken from tasks of a related type may illustrate the influence of accustomed sets. The first is a puzzle employed by Norman Maier (19, p. 142) in which six matches are laid on a table (a plane surface) and the subject is required to construct four equilateral triangles out of them, each triangle having a whole match as a side. In this problem, the habituated set to work in a single plane is reinforced by such indirect suggestions as the plane table, the term "side," and by three dimensional manipulative difficulties. Changing the set to working in three dimensions and using the matches as the *edges* of a regular pyramid having a triangular base, afford the solution. The other example is taken from the New York Times' report (25) of a problem given to some boys

of the Princeton Country Day School, who referred it to Einstein for solution. The problem required two consecutive odd numbers, the difference between whose squares was –56. It further required a statement as to whether or not the numbers were positive *or* negative. The answer was, " 13 and 15, *and* both." Two special conditions should be mentioned here. One is that the 15^2 be substracted from 13^2 or the larger from the smaller number, a procedure contrary to the accustomed set. The other required the recall of a generalization that the square of either negative or positive numbers is positive, although a set to search for a generalization is not habitual in most people. In tasks of these two types, the special sets, the possible generalizations, and other specific conditions, are so numerous that the successful use of such tasks to develop skills in discerning interrelations is not promising. A more desirable type for such a purpose, I consider to be one that assigns the formulation of a generalization, and, at the same time, reveals the relevant conditions in a fairly comprehensive whole, so that the appropriate direction of the learning is inherent in an understanding of the conditions laid down. A serial form that permits constructing many similar problems, all governed by one or a few principles and similar conditions, offers possibilities As to type, the problem reported by John C. Peterson is a good one (32), although a bit too complex. My modification of this problem is described in detail in the Appendix to this article. The specific answers to the first problem are recognizable by almost anyone who has learned to count by twos; but if a " player " is permitted to choose any number from 1 to 10 inclusive, the basic grouping on each successive pair of plays will usually require a generalization for solution. It should be noted that in this task the formulation of a generalization is assigned, and, furthermore, that the non-habituated set to learn how to obtain answers is enforced by the increasing gradation in complexity.

2. *Special Conditions Related to Scoring Possibilities.* An analysis of the learning of a generalization and of kindred discriminative responses reveals certain limitations and confusion with respect to quantitative scoring. It should be evident that

the scoring of the generalization must be of the all-or-none variety; that is, only full success or complete failure is recognized. I attempted at first to utilize part scores, but was forced by various considerations reluctantly to eliminate them. Part scores (in the form of specific memories and answers, promptings and errors, repetitions and times), when critically examined, failed to yield equal scale units for measuring progress toward learning a generalization since certain changes in set may occur which markedly influence performance and yet which are not adequately represented by part-scores. Part scores which were obtained prior to a shift of set, approach, or technique, were manifestly not comparable to those recorded after such a shift. The attempt to formulate part scores in terms of discernment of the interrelations in the complex situations similarly failed. When the recording of relationships less inclusive than a complete formula was requested, poorer students were often satisfied to report principles which would solve only a few problems or to note an observed uniformity, while better students scorned to record them. When the finding of specific answers for problems somewhat graded for complexity offered scorable items based on lower standards of the learner's progress, poorer students at times utilized illegitimate methods and information to get such answers, while better students rejected their use. Furthermore, in order to construct a problem for the utilization of part scores, it was necessary to stress the learning of elements little related to generalization, or to establish lower standards of mastery. Finally, the utilization of part scores also tended to encourage the learner to adopt inefficient modes of attack (such as memorizing and use of specific information), and fostered the unfavorable set to get specific answers, when the assigned goal was to learn how to obtain answers.

This observation of the limitations of part scores in my own tasks suggests that possibly related studies might profit by a critical examination of the objective scores used with reference to the accurate measurement of the discriminative elements involved in achieving the expected goals. The study by Hull (11) on the evolution of concepts suggests the presence of some of the

difficulties which I found. On the other hand, in tasks where the learning of the problem situation is characterized by suddenness, spontaneity, and completeness, attention is apt to be directed to the final step in the learning because it alone is scorable. In such cases the assumption may be made, perhaps unconsciously, that the scorable end-stage is the total learning. This comment may apply to such a generalization as the following: "Learning standard logical relationships is not a slow, gradual process involving practice-effect or trial-and-error responses but occurs suddenly and gives evidence of permanent mastery." (8, p. 291). If such conclusions may be affected by conditions of all-or-none scoring, they need critical examination to reveal the existence or fallaciousness of an underlying assumption that the scorable final aspects are representative descriptions of the *complete* learning process.

3. *Special Conditions Related to the Goal of the Learning.* The presence or absence of knowledge of improvement affects persistence and job satisfactions. When failure is apparent until complete success is achieved, job satisfactions and motivation are likely to be poor. The value which the learner attaches to the achievement of the goal will affect his motivation, and the immediate usefulness of my goals probably approaches zero, if only their intrinsic worth is considered. The objective of using the exercises as a means of developing general facility in doing such tasks is remote, and the skills desired can not be expected to result from learning one generalization; yet immediacy of use has a favorable advantage for good motivation. The reception of these tasks by the subjects indicated poor motivation, especially for the initial exercise where skills usually are undeveloped. My experience has revealed a need for increasing the subject's satisfactions during that period before success is achieved.

4. *Special Conditions Related to 'Distance' Required to Learn the Task.* Different tasks vary in the amount of modification that must be made to achieve the requisite skills or information. A specific learning process involves modification between two points, that point where the learner starts to learn and the goal

of the learning. Considered as a complete task, the discovery of a major premise has length. I have conceived a generalizing problem as modelled along the lines of the requirements for a Ph.D. dissertation. The 'distance' in most generalizing exercises is too great for a task of laboratory length or for ordinary motivation. Some short-cutting is necessary. The short-cutting instituted in the tasks which I have used involved the elimination of the collection of 'data,' and the arrangement of the relevant 'data' so as to facilitate analysis.

5. *Special Conditions Related to Transferability to Other Tasks.* If the learner has developed habitual sets and attitudes which are directionally favorable, and techniques that are efficient, the tasks become relatively smooth running and easy; if, on the other hand, the habitual sets are inhibitory and the skills are inefficient, the tasks become relatively hard and are attended with chaotic frustrations. Assuming that these propositions are valid, the transfer problem becomes one of developing relevant habitual sets and effective skills which are applicable to learning tasks of this type, through the use of suitable exercises and competent instructions. The transferable aspects of generalizing involve responses discriminate in character. They consist of the discernment of influential variables, the discrimination of uniformities in spite of superficial differences, the discernment of how the variables are interrelated and organized and, finally, the formulation of a useful procedure and its verification. The tasks assigned direct attention to these crucial aspects. So analyzed, the learning progress may be expected as a series of discernable "steps" rather than continuous improvement. Initially, my tasks probably give the impression of great difficulty and complexity because of the emphasis on these discriminative aspects alone. However, this artificial isolation of discriminations seems to offer opportunity for perceiving the sets and skills applicable to their learning. The necessity for providing numerous exercises, large amounts of practice, and competent instruction in order to develop transferable sets and skills is, however, not eliminated.

The Skill Components of the Generalizing Task

The skills to be considered are habitual sets, preliminary attitudes, modes of attack, and techniques of manipulation possibly advantageous to making subjective discrimination (28). Initially, the acquisition of such skills is hampered by feelings of chaotic frustration and by the difficulty of measuring discriminative responses. It should be noted that students who succeed, or even prolific makers of generalizations in special fields, fail to show that they can generalize the successful manipulations or discriminative techniques that they use. However, when with several small groups of subjects, an intensive study of as much as six to twelve laboratory hours was devoted to learning such tasks, a marked degree of improvement was conspicuous, as measured by higher percentages of performance in a group. This improvement seems promising for the development of skill components. There was, also, evidence of an increase in job satisfactions in the earlier steps. Since failure is recognized until one begins to formulate the noted uniformities into a useful procedure, sources of satisfaction are small. For the cultivation of such satisfactions, I offer a suggestion derived from the techniques of Edmund Jacobson (12, 13) in teaching "tense" individuals to relax. He first instructs his subjects to contract a given set of muscles maximally, and then to relax them, with a view to teaching them the meaning of muscular hypertension and its reduction. While he stresses the development of *awareness* of the kinesthetic changes, he apparently utilizes the memory of tenseness to develop *satisfactions* from being less tense. Similarly in my tasks, the memories of frustration and of sets that inhibited progress can produce some satisfaction by their very absence when planful manipulations are being used. We may now turn to a consideration of such skill components as habituated sets and techniques.[1]

[1] A greater understanding of the background in which the observations were made may assist in evaluating the material in this section. I first interested myself in generalizing ability, utilizing part scores and six tasks. The performances proved to be poor and part scoring unsatisfactory. Upon modifying the tasks for laboratory exercises, I noted that improvement in ability could be definitely expected with the introduction of attentional factors. When the individual attack was followed by discussion, illustration, and group attack in

1. *The Set to Accept a Generalization as a Goal.* The learner usually expects to get answers, rather than to learn how to get answers. Students often devise non-adaptive goals to which they respond by acquiring more data, checking answers, and manipulating materials beyond all requirements of a clear understanding of the situation.

Let us take, from several possible examples, the familiar bent wire puzzles. The customary task assigned is to learn how to take these puzzles apart and put them together again. It is just as definite an assignment to require the principle of constructing bent wire puzzles so that they can just barely be taken apart and put together again, or to describe the general method of manipulation that will enable the separation of the parts. The accustomed task initiates manipulation; the novel one more likely raises the question, "But what do you want me to do?" Then give them a set of calipers, a finely graduated rule, graduated blade thicknesses, several hardened wire puzzles, and several pieces of soft wire of different diameters and ask specifically, "What are the crucial measurements in the construction of the simple bent nail puzzle?" Students usually do not accept the goal to get the generalization. They want to give you a number such as 1/16 of an inch for the opening, rather than such a generalization as, "The size of the openings that will barely permit separation will be slightly more than one-half the diameter of the wire."

The acceptance of the relational goal will more likely occur if the students have previously learned the desired relationship and only need to recall it or to apply it to a novel situation. In such a case, the learner's effort is directed toward recalling or applying an *old* generalization and away from the set to learn a *new* one. Certainly it is advantageous to accept the experimenter's problem and to cultivate an awareness of the announced goal during all the part activities of the learning.

2. *The Set to Keep all Manipulations Understandable in Terms of How the Parts are Organized in the Whole.* A realization of such an organizing set will be an aid to differentiating the activ-

---

small classes, opportunities for the direct observation of student difficulties, of efficient modes of attack, of the timely introduction of sets, and of the influence of varying directions were afforded. The numerous small sections and need for unfamiliar tasks stimulated the development of additional exercises. Quantitative data were obtained under test conditions from elementary classes, from 523 entering Freshmen, from several classes in experimental psychology, and from individual testing, representing in all about 2,500 papers and 40 different relationships. A brief analysis of certain aspects of the quantitative data will be presented below (*infra*, p. 100). These comments upon skill components and techniques constitute a generalized account based upon experience secured in the extensive experimental program.

ities appropriate to the earlier stages of the generalizing task. The student's past experiences with originals in geometry, with syllogisms, and with algebraic manipulations will have given him a set to start with activities which are applicable, provided some major premise is known. Such a set becomes a barrier to accomplishment when no axiom, relationship, or principle has as yet been learned, since it directs activities toward proceeding accurately from some known principle in accordance with definite rules. Differentiation of the latter set as favorable to the end stages of the task, and an organizational set for the earlier steps will prove advantageous.

Again for our single illustration, let me report observations for the task of finding a short cut method of squaring numbers ending in 5. For this task, I have used the table given below and cardboard pieces to represent units, halves, and quarters, which might be manipulated.

Numbers	Squares
.5	.25
1.5	2.25
2.5	6.25
3.5	12.25
4.5	20.25
5.5	30.25
6.5	42.25
7.5	56.25
8.5	72.25
9.5	90.25

The uniformity of the .25 is readily noted and stimulated the question, "Why?". Also, knowing the relationship of products of the two sides in computing the areas of squares and rectangles, frequently suggests dividing the units of the numbers into the units for the squares. This may suggest noting the uniformity that the quotients are always one more than the divisor. The manipulation of the cardboard halves often suggests utilizing the half units to make whole units. This constructs a rectangle which is always one more than the units on the short side, and always the quarter is left over. After such uniformities have been noted, not before, can the algebraic rules be utilized to aid in securing such a formula as $100X(X+1)+25$, where "X" indicates the number after dropping the first place from the right, or the 5.

3. *The Set to Reduce Complexity by Groupings.* The learner faces frustration often because he seeks to handle too many variables at a time. The complexity of a solid can be markedly reduced if we can deal separately with faces or planes, as can the complexity of areas if we can deal only with linear aspects.

The device of the syllogism keeps the complexity down to two propositions and a relationship, and the derived table reduces it to captions, stubs, and a relationship. The set to combine items in order to make a similar reduction in complexity is an important manipulative procedure, and in such activities the student can derive satisfaction. One method is to substitute a common name for many discrete items.

For example, in the problem for obtaining a short-cut method for squaring numbers ending in 5, naming the units in the table under numbers, "the short side of the reconstructed rectangle" (the 'X' of the formula) and the series of quotients 2, 3, 4, 5, etc., "the long side of the rectangle" ('X+1' of the formula) reduces the task to discriminable dimensions and increases the probability of formulating a generalization.

I have increased by several hundred percent the incidences of success for several tasks in supposedly comparable groups, when I incidentally suggested such combining. This occurred for different revisions of John C. Peterson's Bead Problem, described in modified form in the Appendix, by a mere reference in the description of the task to certain possible choices as the lowest and the highest choices permitted. This set-to-combine reduces the difficulty of discriminating and, in addition, may fortify other favorable organizational sets.

I have given Dr. Joseph Peterson's Rational Learning problem (29) to over 200 adults in the 10 letter form and have analyzed the results for evidence that they had found two favorable methods of organization sufficiently early in the task to eliminate efficiently errors of the logical and perseverative type. In almost all cases, they had failed to leave clear evidence of having utilized both types of favorable organization early in the learning. Modifying the instructions, I suggested the desirability of grouping and stated that a long license plate number of 6 or 9 digits on a moving car could be more easily perceived and remembered if these were considered as 2 or 3 groups of three place numbers each. In analyzing the data of the group to whom was given the grouping set, it was found that both organizational procedures which enabled the elimination of the logical and perseverative errors were adopted at a sufficiently early stage to be definitely helpful. An additional principle was reported by a number of the group receiving the combining instructions, but reported by none of the other group, namely, discarding the digit-letter substitution instructions and learning the problem as a straight position-order task. Furthermore, the level of difficulty was so greatly reduced that the median scores for the 8 and 10 letter forms of this problem were in excess of the 90th percentiles of the non-directed groups for time, repetitions, and all types of errors. Also, the median scores for the much more difficult 16 letter form exceeded the 75th percentile of the easier 10 letter form of the non-directed group. The poorest student in the 16 letter form equalled the 40th percentile

of the 10 letter form group. It should be noted that all subjects taking the 16 letter form had previously had the 10 letter form as a fore-exercise.

4. *The Set to Use Combining and Grouping Techniques as Aids in Discriminating Relationships.* The familiar techniques of the syllogism and derived table are customarily used by instructors to aid them in presenting relationships or by students as an aid in retaining them. By some modification in sets and procedures, such techniques may be employed to facilitate the discovery of relationships. The usefulness of the derived table has become evident in learning several tasks, both in connection with guidance toward desirable sets and procedures, and in test construction. When employed as a device with appropriate sets, it may be used both to aid in discerning uniformities and to assist in the discernment of relationships. The arrangement of the data into this tabular mould directs attention to the discrimination of those variables that are likely to prove influential, since the latter are likely to be selected as stubs for the rows and as captions for the columns. It also directs attention to a systematic arrangement of the data into appropriate cells. With the set to reduce complexity by combining the discrete answers into common categories based on observed uniformities, the systematic arrangement of the data facilities the discernment of such uniformities. With the set to discover relationships between important variables, attention can be directed to the boxes intersecting row and column headings, which in turn means that the learner is working at a level that should admit hope of success. The derived table also enables the subject to utilize attentional factors and principles of learning and of perceptual organization (14, 16, 39, 40). Some of these principles which, with slight modification, are found to be advantageous in the tabular arrangement may be noted:

a. Elements that are in close proximity are likely to be united organizationally. The tabular form gives a visual unity, compactness, and proximity seldom found in other devices.

b. Relationships are more easily discerned in a simple setting than in a complex one. In presenting relationships, comprehension is made easy by a simple situation. In learning relationships,

however, the simple setting gives answers so easily that the relationship is unsought. It is only when answers become difficult to get that a felt need for a relationship develops. The derived table may then present a situation favoring return to the simpler forms. These are usually given in the upper right hand corner of the table—a position relatively advantageous for attention-getting.

A single example of this advantage may be illustrated with the Disc Transfer Problem (31, p. 31). The correct manipulation is expected for a pile of three discs. There is no need to use simpler problems as fore-exercises. Yet a marked rise in percentages of successful generalizations was noted when answers for 1 and 2 discs in a pile were included in the table. Having once discerned that the fundamental uniformity is the successive doubling, the necessary qualification becomes simple only when it is noted that for 1 disc, one moves directly to the appropriate station. The formula 2^n-1 for the series of answers for the number of discs beginning with 1, namely, 1, 3, 7, 15, 31, 63, 127, etc., becomes discernible, where "n" represents both the power and the number of discs in the pile. Yet a successful generalization is difficult to learn without the return to the simplest form.

c. The favorable advantage of the principles of continuity and completeness of data has been well demonstrated in organizing figures presented tachistoscopically. These principles are customarily utilized in the tabular arrangement for a range that permits ready discrimination.

d. The tabular arrangement usually offers a systematization that favors utilizing the principle of inclusiveness and facilitates the perceiving of recurring cycles, rhythms, or various pattern groupings.

e. Another advantage accruing from a table well-filled with answers is that it affords a means of ready, prompt, and accurate checking of uniformities and hypotheses. In several tasks the use of a table has been observed to forestall too narrow and limited generalizations, due to the ease of checking made possible by its use.

f. The tabular device has also enabled the learner who has discerned a similarity, but one which needed revision or qualification to give answers, to hang on to the unworkable principle and to accept the set to revise it in order to make it workable rather than to discard it.

I do not offer the derived table as a royal road to the discovery of relationships, but as one technique which observation of several

tasks has shown to be successful when applied with appropriate sets.

5. *The Set to Try Out Alternative Possibilities.* The tendency to remain in habituated directional ruts has been discussed. It can be avoided by a set to attend to the systematic surveying of possible alternative approaches. Maier has shown that for tasks where inhibitory sets run counter to success, the value of warning the student to avoid habitual directions has increased the number of problems solved from 20 to 40 percent (23, p. 372).

6. *Sets and Techniques to Control Emotional, Affective, and Attitudinal Predispositions.* Most subjects need to recognize the magnitude of emotional handicaps and to adopt sets and techniques to lessen them. In order to overcome them, I suggested to my more intensive study groups the use of the sets and techniques employed by the scientists, jurists, and logicians. For these groups, less evidence of inferiority feelings, and greater disinterestedness, was noted for later than for earlier tasks. The techniques of the scientist might be described briefly as controlled observation, objectivity in the measurement of his observations, quantitative description or measurement, and the use of mathematical, mechanical, and statistical concepts to facilitate grouping and analysis. To these techniques he brings the pose of a disinterested observer seeking to enforce the coöperation of nature by his controlled questionings and observations. The jurist's techniques seek the admission of all relevant evidence that each side brings to his attention and the suspension of judgment until the facts are in, and the similar pose of weighing the facts disinterestedly and impartially without special interest or favor. The logician's technique involves adherence to the rigorous rules of the game in his manipulations of language and symbols, and the same pose of impersonal, impartial disinterestedness in outcomes. I use the term 'pose' advisedly, since poses can be cultivated which lessen the handicaps of felt inferiority and frustration. I include systematic techniques, since it is highly desirable that generalizing be regarded as a form of learning rather than as a vague 'higher' process.

7. *The Influence of 'Time Pressure' Upon Goal Achievement.* It was observed on a few occasions that higher percentages of

successful performance occurred in a group of subjects who were allowed approximately half the time allotted to a presumably comparable group for work on the same tasks. Since the percentages of success were small in both cases and since the observation could have been explained on other grounds, the problem was further studied in a related setting. No verbal formulation of the generalization was required and the checking of specific answers, yielding part scores, was assigned. The question raised was whether or not an impossibly long assignment would tend to enforce the generalizing set and an effort to educe short-cuts, where a more reasonable time limit would enforce the more habitual but less effective sets and methods. Specific answers were required in the following tasks and no verbal formulation of a generalized procedure was assigned.

Tasks were given to two groups of approximately 40 students each. One task was composed of 330 numbers, and two forms were made. The numbers were arranged in numerical order in eleven columns. One form gave numbers ending in 1 starting from 1, the other, those ending in 7. The time limits were 3 and 9 minutes. Both forms were given to each group. One group was given the 3 minute form first, followed by the 9; the other group received the 9 minute form first. Instructions were: "Cross out all non-prime numbers. Your score will be the number crossed out minus three times the wrongs. Use any method, but do not forget that your object is to make a maximum score in the time allowed." I arbitrarily eliminated all records where scores were less than 50 as probably not utilizing short-cuts. I analyzed only scores that met this criterion for the separate forms. Of the group receiving the 3 minute form first, 57% succeeded, to 54% for the group receiving the 9 minute form first, a gain for the time-pressure group of 3%. On the internal evidence of a generalization, the percentage for the group receiving the 3 minute form first is 50; and for the 9 minute form first, 36, a gain of 14% in favor of the 'time-pressure' group.

In order to compare this gain roughly with that dependent upon the giving or withholding knowledge of short-cuts, this task was followed by a second task, which consisted of 330 even numbers of three and four digits each, arranged in a haphazard order, with instructions to cross out numbers divisible by 4 or by 6. They were scored similarly. Employing the same method of comparison, 49% of the group with knowledge of short-cuts and 30% of the group from whom this knowledge was withheld succeeded in meeting the criterion of scores of 50, a gain of 19% attributable to the suggestion to look for short-cuts. The analysis for internal evidence that a generalization was formed gave 46% and 30% respectively, a gain of 16% in favor of the group having knowledge of the existence of possible short-cuts.

These results suggested that both a set to look for generalizations and a feeling of being under the pressure of time, may enforce the set toward relational organization and tend to in-

crease an adoption of new methods which may prove more efficient than habitual ones.

The skill components which have been suggested in this section are in large part attentional factors, attitudinal predispositions, and procedures of manipulation which may be acquired by an individual. Their development is suggested as a way of increasing planful manipulation, of reducing inhibiting sets and procedures, and of aiding the cultivation of job satisfactions in those stages where only failure can be scored over perhaps a long period, even in the cases where successful generalizations are eventually accomplished.

Generalizing Abilities of College Students and Their Correlation With Other Learning Abilities

The incidences of successful performance for unfamiliar generalizing tasks of the type described were low for the groups of untrained Lehigh lower-classmen. Taking fifteen tasks,[2] with no definite instructions concerning economical approaches or techniques of arrangement, re-grouping, or manipulation, the median percentage of successful performance was 11, and the mean percentage was 18.6. In every case, rather narrow limits of time were imposed, since few students working in a group will persist in attacking such problems, continuously recognizing failure, without creating unfavorable group attitudes. No student was permitted to work on a related series of problems longer than 90 minutes. Many were limited to 30 minutes or were given permission to return the task for an easier one, if willing to admit defeat. For three tasks,[3] where additional suggestions were included in the instructions to arrange systematically, to re-group, or to manipulate, the median number of successful performances was increased to 33% and the mean to 24.4%. The number of

[2] See Appendix to this article, problems 1–9 inclusive and 15, 18, 20, 21, 23, and 24. Certain relationship assigned which might possibly be recalled as in portions of 8 and 21 were omitted, and even some of the easier presentations of 10, 14, 17, 19, 26, and 27 were likewise excluded in an effort to secure a rather homogeneous group of tasks and instructions.

[3] See problems 2, 13, and 23. For illustration of the meaning of "additional suggestions" note the final paragraphs in the detailed descriptions of the problems listed in Part I of the Appendix. These are problems 23 and 13.

subjects tested in the above 18 tasks ranged from 590 to 63, for the various tasks. When training was instituted, amounting to from 6 to 12 laboratory hours, and consisting of illustrating the generalized sets and techniques noted in the previous section, and the assignment of additional exercises, the incidence of successes was increased several hundred percent, for several different tasks.

The ability to learn the fifteen tasks without such special suggestions was analyzed for the successful groups on each task, with reference to relationship to the students' general scholastic average (making use of the University's quality point system), and to percentile ratings on the American Council on Education Psychological Examination. The ability to generalize was somewhat positively related both to higher scholarship and to the intelligence test scores. Without presenting the data in detail, the following summarization will indicate something of the degree of the relationships. On the average, about 59 percent of the successful group surpassed the university average scholarship median and about 15 percent surpassed the 90th percentile. None of the students in our 'successful' groups was found to be in the category of students making the lowest passing grade. The relationship of success to scores on the American Council on Education test was higher. About 73 percent surpassed the University median, and about 30 percent surpassed the 90th percentile, while only about 10 percent of the successful group were in the lower quartile. Furthermore, the prediction can be quite confidently made that if a student is not above the lower quartile in either general scholastic average or American Council on Education test score, he will not succeed in learning generalizations of this level of difficulty without additional guidance and suggestions.

Conclusion [4]

This study considers generalizing as a form of learning, examines some special conditions imposed by the type of learning,

[4] At the recent Columbus meeting of the American Psychological Association, Dr. John C. Peterson told me that his brother Joseph had undertaken the preparation of an article on *Insight* and had requested his full notes on the generalizing problem he had published in 1920 (**32**) to support this article. I regard it as a matter of deep regret that death left this desire unconsummated and this contribution unpublished.

develops a general pattern for the construction of exercises to aid in developing skills in generalizing, and illustrates this pattern with detailed examples and references to other tasks employed. It stresses that attention be given to the earlier stages of the learning, to the development of job satisfactions, habitual sets, and techniques appropriate to early stages, and to reducing the length of the tasks to include only the crucial aspects of the generalizing problem. In the support of the observations, inferences, and conclusions given, the experimental evidence is utilized largely in an illustrative manner, due to the wide general scope of the material covered, to the lack of ability of obtaining parts scores for analyzing the skill components, to the difficulty experienced by subjects in reporting their planful manipulations, sets, techniques, or processes for the illumination of their activities, and to the fact that the crucial aspects of the learning are so essentially subjective mental discriminations. The details of experimental findings, therefore, are largely omitted. The limitation of scoring, and the subjective nature of the data available for analysis raise the very fundamental problem of developing new modes of investigating the early stages of the learning in complex settings. In two former studies (6, 7), I have noted the relationship between illusions in very complex settings and the arrest and fixation of the learning process at the early unskilled stages. An experimental attack upon the reasons and conditions which effect such fixations at low learning levels might well throw light upon these obscure and submerged phases.

One quite apparent difficulty is that of selecting tasks for generalizing exercises at the college level. This difficulty is not so great in the lower grades and in animal problems where the stock of simple relationships in simple situations is small but is greatly increased at the post graduate level where generalizing becomes an educational requirement. Adapting the exercises to the student's degree of attainment becomes increasingly difficult as the stock of acquired generalizations increases. Group projects that distribute the drudgery of complete tasks may perhaps abolish the need for short-cutting, increase the motivation to carry through the tasks to completion, reduce the sense of in-

feriority and emotional disruption, and admit of more competent guidance and instruction in connection with such learning.

The advantages of training for skill in generalizing far surpass the hope that important new generalizations will thus be increased. The by-products may be even more valuable in developing attitudes toward the nature and origin of principles, and in releasing learners from the binding effects of inadequate solutions, predispositions, skills, and methods that hold them to low performance levels. Any education is one sided that merely conserves acquired habits, and attitudes. Skills also are needed that equip the individual to adapt successfully to changing cultural demands.

Bibliography

1. Andrews, E. E. An excursion in numbers. *Atlantic Monthly,* 1934, **154**, 459–466.
2. Ball, W. W. R. *Mathematical Recreations and Essays.* London: Macmillan, 1922. Pp. 366.
3. Ballard, P. B. Noegenesis. *J. Gen. Psychol.,* 1929, **2**, 433–443.
4. Boulton, C. L. Nim: A game with a complete mathematical theory. *Annals of Math.,* 1901, **2**, 35–39.
5. Dantzig, T. *Number: The Language of Science.* New York: Macmillan, 1930. Pp. 260.
6. Graham, J. L. Illusory trends in the observation of bar graphs. *J. Exper. Psychol.,* 1937, **20**, 597–608.
7. ——— Illusory trends in the observation of volumes. *J. Exper. Psychol.,* 1937, **20**, 609–619.
8. Hartman, G. W., and Schooley, M., The rôle of insight in the learning of logical relations. *Amer. J. Psychol.,* 1937, **49**, 287–292.
9. Haught, B. F. The interrelation of some higher learning processes. *Psychol. Monog.,* 1922, **30**, No. 139. Pp. 71.
10. Heidbreder, E. An experimental study of thinking. *Arch. Psychol.,* 1924, **11**, No. 73. Pp. 175.
11. Hull, C. L. Quantitative aspects of the evolution of concepts. *Psychol. Monog.,* 1920, **28**, No. 123. Pp. 86.
12. Jacobson, E. *Progressive Relaxation.* Chicago: Univ. Chicago Press, 1929. Pp. 429.
13. ——— *You Must Relax.* New York: McGraw-Hill Co., 1934. Pp. 201.
14. Koffka, K. *Principles of Gestalt Psychology.* New York: Harcourt Brace, 1935. Pp. 625.
15. Köhler, W. *Mentality of Apes.* New York: Harcourt Brace and Co., 1925. Pp. 342.
16. ——— *Gestalt Psychology.* New York: Liveright, 1929. Pp. 403.
17. Lehmer, D. N. Factor table. *Carnegie Inst. Wash.,* 1909, No. 105.
18. MacMahon, P. A. *New Mathematical Pastimes.* Cambridge: The University Press, 1930. Pp. 116.
19. Maier, N. R. F. Reasoning in humans: I. On direction. *J. Comp. Psychol.,* 1930, **10**, 115–143.

20. ——— Reasoning and learning. *Psychol. Rev.,* 1931, **38**, 332–346.
21. ——— An aspect of human reasoning. *Brit. J. Psychol.,* 1933, **24**, 144–155.
22. ——— Reasoning in children. *J. Comp. Psychol.,* 1936, **21**, 357–366.
23. ——— Reasoning in rats and human beings. *Psychol. Rev.,* 1937, **44**, 365–378.
24. Mather, J. E., and Kline, L. W. The psychology of solving puzzle problems. *Ped. Sem.,* 1922, **29**, 269–282.
25. *New York Times,* 1937, June 12 and 17, Editorial pages.
26. Nunn, T. P. *The Teaching of Algebra.* New York: Longmans, 1931. Pp. 616.
27. Peterson, G. M. An empirical study of the ability to generalize. *J. Gen. Psychol.,* 1932, **6**, 90–114.
28. Peterson, J. The effect of attitude on reproduction. *J. Educ. Psychol.,* 1916, **7**, 523–532.
29. ——— Experiments in rational learning. *Psychol. Rev.,* 1918, **25**, 441–467.
30. ——— Intelligence and learning. *Psychol. Rev.,* 1922, **29**, 366–389.
31. ———, and Lanier, L. H. Studies in the comparative abilities of whites and negroes. *Ment. Meas. Monog.,* 1929, **5**. Pp. 156.
32. Peterson, J. C. The higher mental processes in learning. *Psychol. Monog.,* 1920, **27**, 129.
33. Ruger, H. A. The psychology of efficiency. *Arch. Psychol.,* 1910, **19**, No. 15. Pp. 88.
34. Smail, L. L. *College Algebra.* New York: McGraw-Hill, 1931. Pp. 450.
35. Spearman, C. *The Nature of Intelligence and the Principles of Cognition.* London: Macmillan, 1923. Pp. 358.
36. Terman, L. M. *The Measurement of Intelligence.* New York: Houghton Mifflin, 1916. Pp. 362.
37. ———, and Merrill, M. A. *Measuring Intelligence.* New York: Houghton Mifflin, 1937. Pp. 461.
38. Touton, F. C. Solving geometric originals. *T. C. Contrib. to Educ.,* 1924, No. 146. Pp. 115.
39. Wertheimer, M. Untersuchungen zur Lehre von der Gestalt. *Psychol. Forschung.,* 1923, **4**, 301–350.
40. ——— Zu dem Problem der Unterscheidung von Einzelinhalt und Teil. *Zeitsch. f. Psychol. u. Physiol. d. Sinnesorgane,* 1933, **129**, 353–357.

APPENDIX A

A Detailed Description of Two Problems

I. *A Modification of John C. Peterson's Bead Problem* (32).

I. *The Conditions of the Game and the Rules Governing it.* The materials of the game can be any sort of counters (beads, beans, chips, coins). The number of counters placed upon the table at the opening of any game may be any number agreed upon or arbitrarily determined. The two opponents play alternately. A play is made by removing one or more counters, limited to the number permitted, from the original number or from those remaining after preceding plays. The number of counters which may be removed at any play is any arbitrarily determined number. The object of the game is to remove the last counter or to make the final play.

2. *Objective.* (Note that a generalization is assigned.) Describe a method or methods of obtaining "safe combinations," that is, combinations which will enable the player who leaves them to win eventually, even though both players play expertly thereafter.

3. *Safe Combinations.* A 'safe combination' is any number of counters which under the rules of the game will not permit the opponent to win. It can be shown that if player "A" can leave a safe combination for "B," then "B" cannot in turn leave any safe combination for "A," but that "A" on the following play can always leave "B" another safe combination. Hence, as the counters are reduced by successive plays, "A" may always return a safe combination from any safe combination which on the previous play he has left "B," and "B" can never leave such a combination. Thus proceeding, "A" always leaving "B" safe combinations and "B" never able to do so, "A" will eventually win by taking the last counter or making the last play. The generalization problem assigned requires learning how to determine these safe combinations. A list of the safe combinations for a limited selection of choices and for a limited range of open-

ing numbers is given below, in order to make the determination of such safe combinations unnecessary and to make the checking of hypotheses easier.

4. *Interpretation of the Tables.* In the column headings are given all numbers of counters that the players are permitted to

TABLE 1

SHOWING 'SAFE COMBINATIONS' FOR THE RANGES OF CHOICES INDICATED IN THE COLUMN HEADINGS (CONTINUOUS SERIES)

Permitted Choices

1	1–2	1–3	1–4	1–5	1–6	1–7	1–8	1–9
2	3	4	5	6	7	8	9	10
4	6	8	10	12	14	16	18	20
6	9	12	15	18	21	24	27	30
8	12	16	20	24	28	32	36	40
10	15	20	25	30	35	40	45	50
12	18	24	30	36	42	48	54	60
14	21	28	35	42	49	56	63	70
16	24	32	40	48	56	64	72	80
18	27	36	45	54	63	72	81	90

Permitted Choices

2–3	2–4	2–5	2–6	2–7	3–4	3–5	3–6	4–6
5	6	7	8	9	7	8	9	10
6	7	8	9	10	8	9	10	11
10	12	14	16	18	9	10	11	12
11	13	15	17	19	14	16	18	13
15	18	21	24	27	15	17	19	20
16	19	22	25	28	16	18	20	21
20	24	28	32	36	21	24	27	22
21	25	29	33	37	22	25	28	23
25	30	35	40	45	23	26	29	30

remove at a given play. In the column below each heading are given all safe combinations for these choices within the limited range of the smallest nine arranged in consecutive order from the lowest opening number of counters. In reading the column headings, or choices permitted, the hyphen indicates that all numbers inclusive of the first and last may be removed; and the comma indicates that only the numbers indicated are permitted. For example, '1–6' means that 1, 2, 3, 4, 5, or 6 counters may be removed at any play, but '1, 6' means that only 1 or 6 may be removed.

All the answers in the first block and some of those in the second may be secured by the learning of a simple relationship.

The additional answers in the second block may be determined by a supplementary relationship, or may be stated as a qualification to the fundamental principle. Describe the method that will enable you to determine these safe combinations. These relationships should be sufficient to enable determination of the safe combinations for all those of similar continuous series, that is, all situations in which any number may be removed at any play between and inclusive of the highest and lowest choices permitted.

The next two blocks of answers might be designated as a discontinuance series, in that the intermediate numbers are not permitted. Either an additional principle is needed for them or a still broader or more inclusive statement of the fundamental principle is required. Note that a statement of principles that is adequate, but which is not the most inclusive statement of the

TABLE 2

Discontinuous Series: Permitted Choices Are Indicated in the Column Headings, and the Corresponding 'Safe Combinations' Are Shown in the Columns Below

Permitted Choices							
1, 3	1, 5	1, 7	1, 9	1, 4	1, 6	1, 8	1, 10
2	2	2	2	2	2	2	2
4	4	4	4	5	4	4	4
6	6	6	6	7	7	6	6
8	8	8	8	10	9	9	8
10	10	10	10	12	11	11	11
12	12	12	12	15	14	13	13
14	14	14	14	17	16	15	15
16	16	16	16	20	18	18	17
18	18	18	18	22	21	20	19

Permitted Choices									
2, 6	2, 10	3, 9	3, 15	4, 12	2, 4	2, 5	3, 7	4, 9	3, 5
4	4	6	6	8	6	4	6	8	8
5	5	7	7	9	7	7	10	13	9
8	8	8	8	10	12	8	11	14	10
9	9	12	12	11	13	11	12	15	16
12	12	13	13	16	18	14	16	16	17
13	13	14	14	17	*19*	15	*20*	21	*18*
16	16	18	18	18	(4)	*18*	(7)	*26*	(6)
17	17	19	19	19	(5)	(5)	(8)	(9)	(7)
20	20	20	20	24	(10)	(12)	(17)	(10)	(14)
21	21	24	24	25	(11)	(19)	(18)	(11)	(15)

fundamental relationship, will without qualification give the losing combinations inclosed in parentheses. See if the relationships can be so qualified or stated that these losing combinations will be eliminated.

4. *Solution.* The fundamental relationship pertains to understanding the organizational possibilities for a single pair of alternate plays. Such safe combinations may be expressed as the lowest plus the highest permitted choices, or, for an added principle in the discontinuance series, as twice the lowest choice. Multiples of these numbers will also be safe combinations, since by grouping one may, also, always divide the original number of counters into groupings of the simpler problems. Due to the fact that the final play is also a winner, any combination that leaves less than the lowest permissable choice becomes a safe combination. Hence, a supplementary principle that 1, 2, 3, 4, 5, up to the lowest choice permitted minus 1 should be added. Qualifications of the fundamental principle of the lowest plus the highest choices permitted are not required in either series. However, multiples of twice the low are nullified by the intermediate choices in the continuous series. In the discontinuous series, the multiples of twice the low are only applicable if they are less than the highest choice. The supplementary principle of 1, 2, 3, etc., up to the lowest choice minus 1, when applied to twice the low multiples is similarly modified by the high choice.

This problem was added to the group of "additional suggestions" to take hold to manipulate planfully, by suggesting a survey of the "safe combinations" and all possible manipulations involving opening numbers not greater than the highest that could possibly be reduced by a single pair of alternate plays for B to win.

II. *Problems Involving the Use of Factor Stencils* (17).

Objectives. Problem 1: Describe a method of applying short factor stencils to a long table to obtain all numbers evenly divisible by the specific factor of the key stencil. State the relationship in such general terms that the derived procedure is sufficient for

any prime factor stencil. Problem 2: Describe a method of utilizing prime factor stencil keys to obtain non-prime factor stencil keys without additional stencil cutting. Note that the prime stencil keys may require added length to give all answers. The following supplementary principle was added: Describe in general terms the minimum length of the key required to give all correct answers. Problem 3: Describe a method which will tell what stencils, other than prime stencils, must be cut to obtain stencil keys for all possible factors.

1. *Materials Provided.* 1. A set of perforated stencil keys for the earlier prime number factors, namely, 2, 3, 5, 7, and 11. 2. A tabular arrangement to identify consecutive numbers up to 510 for use with the short prime stencils. 3. Directions for the application of the short stencil keys for the top of the table, namely, for numbers 1–99 inclusive. 4. Added suggestions for manipulating in a different context for the two later problems.

2. *Interpretational Comments.* 1. A "factor stencil" is a stencil showing perforations, which, when properly applied to the tabular arrangement for identifying consecutive numbers, will indicate those numbers and only those numbers that are evenly divisible by the factor for that stencil. 2. A "prime factor" is a prime number, that is, a number which can only be divided by itself and 1. 3. A "relationship" is a description of a connection, which, in this case bears upon the question of how to proceed to get answers. 4. The method of using the table for identifying consecutive numbers is as follows: The numbers heading the column represent the units of any number. Note that these numbers also appear on the stencil keys. The numbers heading the rows are consecutive multiples of ten. Any number up to 510 can be identified by noting the position of lines intersecting row and column headings and adding the units of the column headings to the number of the corresponding row. Thus the number 39, the sum of 30 and 9, is found at the intersection of lines for the column headed 9 and the row headed 30. Similarly 130 is found at the intersection of the row headed 130 and the column headed 0.

3. *Manipulative Procedures and Suggestive Arrangements.* (a) For Problem 1: Place each stencil provided upon the table so that the zero of the top row of the stencil is adjusted upon the zero of the table. Note that the perforations give correct answers for the factor used, up to 99. Learn to adjust the stencil upon the various parts of the table so that correct answers and only correct answers are given. Seek to formulate in general terms instructions for adjusting these stencils so that they will be equally applicable for the adjustment of any of the five stencils provided. Note that for prime stencils higher than 11, longer stencils must be utilized.

(b) For Problem 2: If the stencils and table are properly adjusted for length, the following additional factor stencils may be made, using only the five stencils provided: 6, 10, 14, 15, 21, 22, 30, 33, 35, 42, 55, 66, 70, 77, 105, 110, 154, 165, 210, 231, 330, 385, 462, 770, 1155, 2310. Learn to construct such stencils and to formulate the derived method so that one set of instructions will apply to the various factors given above. Starting with the lowest factor in the list, search for a method of using the prime stencils to obtain a stencil for the new factor, then try it for the next factor higher, etc. Test your method to determine whether or not it is stated broadly enough to apply it to the factors for 30, 42, 66, 70, and higher numbered factors. If you fail to discover a method of constructing such stencils, it is suggested that you utilize the different context and procedure given in the final paragraph of Problem 3.

(c) For Problem 3: In order to provide factor stencil keys for all factors, 2–100 inclusive, additional prime factor stencils must be cut, namely, 13, 17, 19, 23, 29, 31, 37, 41, 43, 53, 59, 61, 71, 73, 79, 83, 89, and 97, but only the following non-prime factor stencils, namely, 4, 8, 9, 16, 25, 27, 32, 49, 64, and 81 need be cut. Learn how to determine just what non-prime stencils must be cut and to formulate a generalization to describe the derived procedure.

The following context and procedure may assist you in seeking the desired relationship. Place in a row the consecutive numbers

beginning with 1, as in the arrangement below. Then write below each number the lowest prime factor which will factor that number, then below that, the lowest prime factor that will factor the quotient obtained, etc., until the numbers may be represented as the products of the prime factors. Study the arrangement for uniformities existing in the correct answers.

1,	2,	3,	*4,*	5,	6,	7,	*8,*	*9,*	10,	11,	12,	13,	14,	15,	*16,*	17,	18,	19,	20,	21,	22,	23,	24,	*25,*	etc.
			2		2		2	3	2		2		2	3	2		2		2	3	2		2	5	
			2		3		2	3	5		2		7	5	2		3		2	7	11		2	5	
							2				3				2		3		5				2		
															2								3		

APPENDIX B

A Classified List of Exercises for Developing Generalizing Ability

The difficulty of finding suitable tasks for the college and superior adult level has been commented upon. I am, therefore, including a list of the tasks which I have used, or which, in a few cases, are similar to those used. The tasks vary in inherent difficulty and complexity, and with method of presentation.

I. *Tasks involving summation series:* The derivation of a *method* of computing answers to the following problems is required, and appropriate illustrative answers for a limited range of each series is stated.

1. Using a series of equilateral triangles with the area uniformly dotted and the triangle built up with 1,2; 1,2,3; 1,2,3,4; etc., dots in a row, a method of computing the number of dots in the triangle from knowledge of the number of dots on a side is required (5, pp. 41–43).

2. Using a series of squares with 2,3,4,5, etc., dots in rows and columns, each square being divided into two right triangles, a larger one employing the greatest diagonal, the smaller one, the adjacent exterior diagonal, a method of computing the number of dots in the large and, also, in the small triangle, from knowledge of the number of dots on the side of the square, is required (5, pp. 41–43).

3. Using the faces of regular solids for figures, the method of computing the total dots in the sides of the figure, from knowledge of the number of dots on any side is required. Similar tasks are assigned, involving the edges of the regular solids, and areal tasks, involving one or all faces of the regular solids.

4. Using series of pyramids having equilateral triangle or square bases, and marbles in the place of dots, tasks similar to the above are given, requiring a method of computing the number of marbles in a pyramid from knowledge of the number of marbles on the edge of the pyramid.

5. Using the construct of Pascal's triangle, and the terms of

the series in any column, the subject is asked to develop a method of getting answers for the various columns' series and a general formula for all series from knowledge of the number of the term. A supplementary principle is required for a method of constructing this triangle (34, p. 180).

6. Using a series of squares ruled off into one centimeter units of 1, 2, 3, 4, etc., centimeters in a row or column, a method of computing the number of different squares possible from knowledge of the number of centimeters in any row is required. Similar tasks are set for the number of different rectangles, remembering that a square is always a rectangle but that a rectangle is not always a square; and for a cube, to determine the number of different cubes in a cube, and the number of rectangular solids in a given cube.

II. *Tasks involving difference series:*

7. A series of tasks is assigned, utilizing interpolation tables and supplementary tables developing the method of constant differences. The tables give only integer values, and a method of interpolation to secure accurate answers for decimal values is required.

III. *Tasks involving product series:* In general, from knowledge of a number in an arithmetical series and appropriate answers for a limited range, learning a method of computing answers is required.

8. Using Atkinson's Ingenuity Test situation, the method of computing the number of possible combinations from knowledge of the number of letters in any row is required. A systematic method of arranging answers to eliminate duplication is also required. An extension of this task added all possible duplications of the correct responses. Other variants are "Kirkman's Problem" and bridge arrangement problems (2, p. 237; 31, pp. 114–115; 2, Ch. 10).

9. Using a series of city maps of uniform block design, a method of computing the maximum number of ways, as short as the shortest, of traversing these square and rectangular maps from the NW to the SE corners from knowledge of the number of blocks in rows and columns is required.

10. Using levers of the first class placed in equilibrium (27),

or the correlation array with the mean lines drawn in, the problem is developed for the re-discovery of the principle of moments.

11. Using a series of domino sets of the ordinary rectangle of two squares, or the equilateral triangle with three similar triangles, the generalization of the number of dominoes from knowledge of the number of dots permitted is required (**18**).

12. An extension of the problem of generalized dominoes, called "Designing Pastimes," involves introducing arbitrary restrictions of patterns or boundaries (**2**, pp. 64, and 255; **18**).

13. A factor stencil problem has been described in detail in Appendix A (**17**).

14. The re-discovery of Eratosthenes Sieve for sorting out prime numbers has also been assigned (**2**, **17**).

IV. *Tasks involving progression or power series:*

15. Using the Disc Transfer Puzzle (**2**, p. 288; **31**, pp. 31–32), a method of finding the fewest possible number of moves from knowledge of the number of discs in a pile is required. A supplementary principle describing a method of moving accurately is also required.

16. The Chinese Ring Puzzle supplies material for a similar set of generalizations, and possibly other puzzles of Ruger's list could be readily adapted for the construction of similar series (**2**, **33**).

V. *Other tasks essentially mathematical in type:*

17. A short cut method of squaring numbers ending in five has been described in the text of this article.

18. Other short-cut methods are usable, such as shifting from the decimal to the duo-decimal system of notation (**1**), or from one of several systems of measurement to another.

19. The water-dipping ingenuity test has received mention in the text (**36**, **37**). See also "Bachet's weight problems" (**2**, p. 34).

20. The checker puzzle problem (**9**, p. 34; **2**, pp. 77–80) provides a generalization for the number of moves and jumps from knowledge of the number of checkers used, and one for checking accuracy in moving and jumping.

21. Employing the familiar tabular device and utilizing an arithmetical series for the stubs of the rows, the relationship

between the rows and a column of answers, affords many possible generalizations, such as reciprocals, $N^2 — N$, etc.

22. The number series device in appropriate sets is usable.

VI. *Games involving uniformities and complete solutions:*

23. A series based upon John C. Peterson's study has been described in detail in Appendix A (**32**).

24. The Game of Nim (**4**) gives the initial impression of extreme difficulty. In this game, two or three piles of counters are placed upon a table. Two opponents play alternately and are permitted to remove one or more counters from one and only one pile at any play. The object of the game is to leave safe combinations for the opponent in order to win by removing the last counter. A rule is sought for determining how to compute the safe combinations.

VII. *Miscellaneous tasks:*

25. A bent wire puzzle series is given brief mention in the text (*supra,* p. 93).

26. Using artificial language constructs and giving vocabularies and translations, the learning of the grammatical and syntactical rules is required.

27. Using a series of figures for continuous tracing without raising the pencil, crossing, or re-tracing lines, a generalization may be required which will tell what kind of figures are unicursal, which are not, or where to begin and end such figures (**2**, Ch. 9; **9**, p. 37). Variants include the puzzle of moving pawns along straight lines from some vacant corner to fill one less than the number of points; and the generalized order of travel that enables one to visit once and only once each point of a plane projection of each of the regular solids.

28. A series of chess or peg removal problems involving jumping and removing pieces may be constructed and the appropriate generalized procedures required (**2**, p. 109 ff.).

29. A series of block moving puzzles involving space restrictions and specific start and end patterns is similarly usable (**2**, p. 224).

30. A series of plane match puzzle patterns involving specific beginning and end patterns is likewise usable.

Reprinted from the JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
Vol. 23, No. 1, July, 1938
Printed in U. S. A.

AN EXPERIMENT IN GENERALIZING: A UNICURSAL PROBLEM (1)

BY JAMES L. GRAHAM

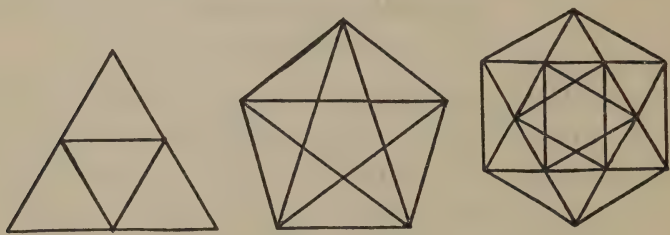
Lehigh University

The problem was derived from the familiar puzzles of tracing geometric figures without lifting the pencil from the paper, crossing, or retracing lines. The task assigned was the discovery of generalizations of the major premise type and not the ability to trace unicursally the forty-eight figures presented. Success typically involved not a single but a series of discriminative acts. Three good generalizations and a poorer one were discovered. These are: 1. Any figure which has no odd branchings at any point can be traced unicursally, starting at any point. 2. Any figure which has two and only two odd branchings can be traced unicursally, if the start and end occur at the odd branchings. 3. Any figure which has more than two odd branchings can not be traced unicursally. 4. The poorer principle described a method of tracing figures and contained two or three elements, namely, a method of grouping the complex figures into *n* unicursal paths, a method of proceeding sequentially from one such path to another adjacent unicursal path, and a method of avoiding the improper closing of the adjacent unicursal path. There were greater difficulties in stating the fourth principle in universal terms and in expressing it clearly. Since it was scored more liberally and was closer to the more habituated task of discerning one's ability to trace the complex figures, it has been analyzed separately.

The procedure used follows: A fore-exercise was presented orally to illustrate four types of discriminative acts logically involved in the discovery of such generalizations. It discussed quite a different problem from the unicursal one, namely, the discovery of a shortcut method of squaring numbers ending in five. The unicursal problem was presented

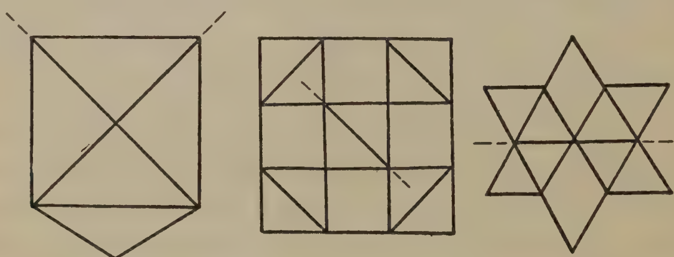
in written form. It described the familiar puzzle situation, and presented and partially analyzed forty-eight figures grouped into a set that could be traced unicursally beginning at any point, a set that required starting at designated points,

FIG. I. Such figures can be traced unicursally, starting at any point.



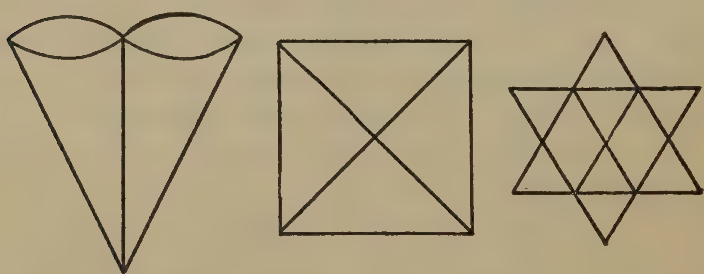
Lines 9	Lines 20	Lines 42
Points 6	Points 10	Points 18
Branchings 3(2's) 3(4's)	Branchings 10(4's)	Branchings 12(4's) 6(6's)

FIG. II. Such figures can be traced unicursally, if start is made at a dotted line.



Lines 10	Lines 25	Lines 22
Points 6	Points 16	Points 13
Branchings 1(2) 2(3) 3(4)	Branchings 4(2) 10(4) 2(5)	Branchings 6(2) 4(4) 2(5) 1(6)

FIG. III. Such figures cannot be traced unicursally.



Lines 7	Lines 8	Lines 22
Points 4	Points 5	Points 13
Branchings 3(3) 1(5)	Branchings 4(3) 1(4)	Branchings 6(2) 3(4) 4(5)

and a set that could not be traced unicursally. The students worked on the task for thirty minutes in groups of 38 to 40. The four types of discriminations pointed out in the fore-exercise, with an illustration taken from the unicursal problem instead of from the fore-exercise, are given below: 1. The discernment of crucially important variables from those that are inconsequential or irrelevant. It seems obvious that to succeed the students must discriminate the crucial influence of the order of branchings from any point, from that of the number of points, lines, etc. 2. The discrimination of uniformities amid diverse specific elements. This is essentially a process of grouping to reduce the complexity of the task to discriminative levels. On the page of figures which could be traced unicursally, starting at any point, the order of the branchings from any point were, 2, 4, 6, or 8. These could be grouped under the common description of 'even branchings' at every point. 3. The discernment of relationships. It may be noted that with an even number of branchings at a point, each pair provided an entrance and an exit from that point. 4. The discernment of a formulation useful from the standpoint of applying the previous discriminations to the characteristic goals desired. The four generalizations listed are examples of this final discriminative act. An experimental analysis of these discriminations and their sequential order in discovering relationships, is hampered by the fact that these discriminations are not regularly available as part scores and are not meaningful as separable generalizations until they culminate in the formulation of a useful procedure relative to the desired goal. Students frequently fail to verbalize these discriminations and even when successful fail to describe the steps through which they achieved success. They customarily have difficulty in discerning them as places to take hold planfully and know no place to attack fruitfully. In fifteen other problems constructed along similar lines, this fact has been shown. When no suggestions were given as to the types of discrimination involved, the percentages of success were a third to a fifth less than in this task where they were illustrated in a fore-exercise (**2**). Results: Fifty-five percent of 152

students learned one or more of the good generalizations listed above, and seventy-seven percent succeeded if all four principles are included. As a group, they averaged better than one 'good' generalization apiece, and 1.35 per student when all four principles are included. Correlation with the

TABLE I

Comparison of Generalizations and Scores on the A.C.E. Examination

Generalizations	A.C.E. Quartiles				A.C.E. Deciles	
	Q_1	Q_2	Q_3	Q_4	1	10
	Percentages of individuals succeeding					
1, 2, and/or 3	17	23	27	32	2	16
4 only	38	31	15	15	31	7
None	26	25	32	17	10	0
None and 4	33	28	23	17	20	3
	Percentages of generalizations offered					
1, 2, and/or 3	19	20	27	34	2	17
4 only	33	32	15	20	23	7
1, 2, 3, and/or 4	22	23	24	31	6	16

TABLE II

Comparison of Generalizations and Psychology Grades

	Grade Quartiles				Grade Deciles	
	Q_1	Q_2	Q_3	Q_4	1	10
	Percentages of individuals succeeding					
1, 2, and/or 3	21	22	24	33	7	13
4 only	29	31	27	13	9	5
None	28	28	26	18	17	8
None and 4	29	28	27	16	13	7
	Percentages of generalizations offered					
1, 2, and/or 3	21	20	23	36	7	15
4 only	29	24	32	15	9	6
1, 2, 3, and/or 4	23	25	25	28	8	13

The generalization numbers are those used in the text of the article. The numbers in the table represent distribution percentages not absolute scores, and these have all been equated for the size of the groups.

A.C.E. Examination was + .43 using only the first three principles, and + .37 using all four. Correlation with psychology grades was + .27 for three principles, and + .33 for four. Critical levels are probably more significant. Thus a comparison of the highest and lowest deciles, after equating for the size of the groups, shows eight times as many generalizations and successful individuals in the highest decile of the A.C.E. Examination, but only twice as many in the highest decile in grades.

Motivational factors present a serious difficulty in these tasks, since failure is recognized by the student until a useful formulation has been completely acquired. Students are observed to show marked individual differences in the time and effort expended before they reject the assignment or substitute for the generalizing task one of getting specific answers such as learning to trace particular figures unicursally. The difficulty of finding suitable part scores for analyzing the amount of frustration a student will bear before quitting or diverting his interest from the task, hampers an attack upon this problem. However, the situation here is much simplified over the usual adjustment and personality situations, where this problem assumes practical importance. A similar absence of meaningful and constant part scores, since only full success or complete failure is scorable, hinders the analysis of the higher learning processes involved in the discovery of generalizations of this major type. However, the isolation of tasks that are essentially discriminative from other learning types should have some possibilities for fruitful research. Our educational programs and procedures have too long neglected the development of skills in discriminating and generalizing, through an over emphasis upon the application of generalizations as a means of securing specific answers. The improvement of guidance methods and knowledge of the prerequisite skills are, therefore, most desirable.

REFERENCES

1. Ball, W. W. R., *Mathematical Recreations*, Macmillan and Co., Ltd., 1922, pp. 170–181.
2. Graham, J. L., Learning to generalize, *Psychol. Monog.*, 1938. (In Press.)

(Manuscript received February 8, 1938)

LEHIGH UNIVERSITY PUBLICATION

(List continued from inside front cover)

* Limited number of copies available for free distribution.

LEHIGH UNIVERSITY STUDIES

Psychology Series, No. 2

Evidence Against the Alrutz Theory of Heat

by

WILLIAM LEROY JENKINS
Assistant Professor of Psychology

Reprinted from the Journal of Experimental Psychology

LEHIGH UNIVERSITY
BETHLEHEM, PENNSYLVANIA

1939

Lehigh University Studies comprise selected papers arising from the creative work of the University which for any reason are not issued as periodical bulletins from the **Institute of Research.**

LEHIGH UNIVERSITY STUDIES

Science and Technology Series

No. Title

* 1. Tables of the Properties of Aqua-Ammonia Solutions. 15 cents.

Psychology Series

No. Title

* 1. Learning to Generalize and An Experiment in Generalizing a University Problem. 15 cents.

* 2. Evidence Against the Alrutz Theory of Heat. 15 cents.

---

* Limited number of copies available for free distribution.

Reprinted from the Journal of Experimental Psychology,
Vol. 22, No. 6, June, 1938
Printed in U. S. A.

STUDIES IN THERMAL SENSITIVITY: 6. THE REACTIONS OF UNTRAINED SUBJECTS TO SIMULTANEOUS WARM + COLD + ELECTRIC SHOCK

STUDIES IN THERMAL SENSITIVITY: 6. THE REACTIONS OF UNTRAINED SUBJECTS TO SIMULTANEOUS WARM + COLD + ELECTRIC SHOCK

BY WILLIAM LEROY JENKINS

Lehigh University

In a previous article,[1] the reactions of untrained subjects to simultaneous warm + cold stimulation were described. Unlike the traditional results with trained observers, these untrained individuals did not report 'hot' from compound grill stimulation, except in a few scattering instances. In fact, the intensity of the warmth seemed to be diminished rather than accentuated by the addition of cold. To account for the discrepancy, it was suggested that the common 'hot' may be a different experience from the trained observer's 'heat'; and that the Alrutz theory may apply to the latter, but not to the former.

The synthesis of 'burning heat' from warm + cold + pain has also been cited as evidence for the Alrutz theory. Ferrall and Dallenbach [2] secured the addition of pain to the warm-cold complex by using a low-voltage alternating current wired to alternate tubes of the grill. The temperature of the cold system was varied from 9° to 20° C., and of the warm system from 38° to 42° C. (for observer *J*) and from 40° to 44° C. (observers *A* and *D*), to keep below the heat limen previously determined for each individual. The electric current was varied between 10 and 15 volts at random. Control tests showed that 'burning heat' was not reported from the electric shock alone when the grill systems were at skin neutral. With

[1] Jenkins, W. L., Studies in thermal sensitivity: 5. The reactions of untrained subjects to simultaneous warm + cold stimulation, *J. Exper. Psychol.*, 1938, **22**, 451–461.

[2] Ferrall, S. C. and Dallenbach, K. M., The analysis and synthesis of burning heat, *Amer. J. Psychol.*, 1930, **42**, 72–82.

warm + cold + shock, however, burning heat was reported by the three observers as follows:

D (trained in introspection and experienced in grill experiments)—in 34 out of 40 trials.
J (trained in introspection but with no previous experience in cutaneous experiments)—in 12 out of 36 trials.
A ('totally untrained in introspection')—in 6 out of 48 trials.[3]

Concerning the discrepancy, the authors say: "The difference between the results of *D*, on one hand, and *A* and *J*, on the other, may be due to:

(1) 'set'; *D* knew the object and plan of the experiment, and he may unconsciously have been biased in his reports in the direction of burning heat (*D* admits this possibility, but he is certain that he had not anticipated a particular result, and that the qualities he reported were matters of indifference to him);
(2) difference in experience and training. *D* was highly trained in cutaneous observation in general and in the introspection of heat in particular; *J*, though an experienced *O*, had never before observed in a cutaneous experiment; and *A* was a novice;
(3) differences in the effectiveness of the stimuli—particularly the electrical stimuli. The cutaneous stimuli were, as the experiments of Series V indicate, more effective for *D* than for either *A* or *J*, possibly because of differences in the amount of pressure of the arm upon the grill or in the amount of perspiration which bathed the arm."[4]

Some years earlier, Knight had made a study of the integration of punctiform warm + pain.[5] Mapped 'warm spots' were stimulated by radiant means until a moderate experience of warmth ensued. Then the skin nearby was lightly pricked

[3] *Op. cit.*, 79–80.
[4] *Op. cit.*, 80.
[5] Knight, L., The integration of warmth and pain, *Amer. J. Psychol.*, 1922, **33**, 587–590.

with a needle. Three introspectively-trained observers were used. The results are not reported quantitatively, but the first sentence of her conclusions is most suggestive: "The simultaneous stimulation of warm and pain spots may result in an experience which is variously called 'burning heat,' 'burn' and 'hotness.'"[6]

The significance of these findings seems to have been generally overlooked. Taken literally, they mean that burning heat can be synthesized from warm + pain, with no cold present. If we take 'hotness' to be the equivalent of 'heat,' then heat itself can be experienced with just warm + pain. In either case, the implication is that cold is not a necessary condition for synthetic heat.

Unfortunately, the integration of warm + pain in the grill technique does not seem to have been tried with trained observers. Since only untrained subjects were at our disposal, we determined to attempt the experiment with them. If any positive effects were obtained, it would be possible to compare the results from warm + cold + pain with those from warm + pain. If the original simple stimulation were reported as 'hot,' it would be possible to see if this could be changed to 'burning hot' by the addition of simple electric shock.

Apparatus and Procedure

Grill No. 1 from the warm + cold experiments was used throughout. In this grill, the cold bars can be dropped about 5 mm below the surface and raised into position by a cam action. Warm and cold bars are 3 mm wide, with approximately a 1 mm air space. Alternate bars were wired to the two leads from a transformer connected to the lighting circuit. The current used was approximately 9 volts.

Since an important variable in shocking is the condition of the skin, the surface of the grill was moistened with a strong salt solution before each stimulation. In spite of this precaution, it soon became evident that the effectiveness of the shock was not the same for all subjects. Some experienced a sharp pain; others only a mild tingling; while in a few cases there was apparently no effect at all.[7] Control tests give some idea of the general effectiveness of the shock. With both temperature systems at neutral, the shock alone produced some non-thermal report (tingle, sting, ache, pain, etc.) in 83 percent of the subjects. About 10 percent of them, moreover, reported the simple shock as 'hot' or 'burning hot.'

---

[6] *Op. cit.*, 590.

[7] Ideally, the strength of the current should be adapted for each individual. Practically, it was impossible to do this without disclosing the technique of the experiment to the subjects.

The procedure was the same with all subjects. The simple warm stimulation was given first. After five seconds, the movable bars were brought up in contact with the skin and the shocking current turned on. After five seconds more, the subject removed his arm and went out to write his report. Subjects were handled in groups of about ten, and with a few exceptions had also served in the earlier experiments upon warm + cold stimulation. They were not informed of the plan or purpose of the experiments, which most of them seemed to assume were merely a continuation of the earlier work.

All reports were made on the same form, which had separate columns for Experience *A* (first five seconds) and Experience *B* (second five seconds). The formal instructions were:

"In the columns headed 'Therm.' report all thermal (temperature) sensations in each experience, using the following symbols:

V. C.—very cold	O—neutral	V. W.—very warm
Cold—cold	S. W.—slightly warm	Hot—hot
Cool—cool	Warm—warm	Burn—burning hot

In the columns headed 'Non-therm.' report all non-thermal experiences, particularly pain. Try to characterize each experience with a single word if possible, such as 'tingle,' 'prick,' 'sting,' 'ache,' etc."

Possible Artifacts

In addition to the possibility that electric shock alone may be reported rather rarely as 'hot' or 'burning hot,' there are two possible sources of artifacts in the experiments as conducted.

1. All temperatures were measured at the *inlet* and are not necessarily the temperatures of the grill surfaces themselves. Calibration tests showed that the outlet temperature of the warm water was not appreciably affected by the temperature of the water which flowed through the cold bars; whether these were in the raised or dropped position, the inlet-outlet differential remained about 2° C. But this is no guarantee that the surface temperatures were not altered. Since multiple thermocouple-galvanometer measurements of the surface temperatures were not feasible, a physiological check was substituted. If the surface temperature of the warm bars was materially affected by the temperature of the intervening cold bars, this should show in a reduced percentage of 'hot' and 'burning hot' reports with simple stimulation (*i.e.*, with the cold bars in the dropped position). Table 1 gives the results of this

TABLE 1

Percentage of Hot and Burning Hot Reports with Cold Bars in Dropped Position

Cold Bars at	Warm Inlet Temperatures					
	34°	40°	42°	44°	46°	48°
34°	0.0	5.1	10.6	19.6	59.9	84.2
25°	—	1.2	10.8	21.5	48.5	—
20°	—	5.6	7.4	32.6	54.7	—
15°	—	5.3	10.6	22.9	39.5	—
10°	—	3.3	5.0	20.0	17.5	—

check. With cold temperatures of 25° and 20° C., no effect is evidenced. With cold temperatures of 15° and 10° C., there may be some, especially with the higher levels of warmth. This should be taken into account in interpreting the results of 10° and 15° C. compound stimulation.

2. Using the same *order* of stimulation throughout may have introduced another artifact. Even without shock, when the arm is in continuous contact with the grill, the experience may increase in intensity so that the report for the second five seconds might naturally be higher. On the other hand, adaptation might occur so quickly that the intensity of the second report would naturally be lower. A partial check is available from the effects of 48° and 50° C. stimulation without shock (see Table 2), designed

TABLE 2

COMPARISON OF EXPERIENCES *A* AND *B* WITHOUT SHOCK

Change	48°	50°
Same report	51.6%	59.2%
Warm to hot	16.5	7.4
Warm to burning hot	2.2	12.4
Hot to burning hot	16.5	14.8
0 to warm or hot	0.0	2.5
Decreases (hot or burning hot to warm)	13.2	3.7

primarily to show that subjects would report 'hot' and 'burning hot' with simple stimulation at high temperatures.[8] At 48° C., the *A* and *B* reports were the same in somewhat more than half the cases, but with the balance, there is a definite preponderance of increases over decreases. It is regretted that data are not available for the lower temperatures.

Results

The figures presented in Table 3 are percentages of 'clear chances.' There was a 'clear chance' for a change from warm → hot or warm → burning hot, only when the original report was not 'hot' or 'burning hot.' Similarly, there was a 'clear chance' for a change from hot → burning hot only when the original report with simple stimulation was 'hot.'

1. With the movable bars at 34° C., the addition of *shock* results rather uniformly in an increase in the intensity of the experience reported. An original report of 'warm' is changed to 'hot' or 'burning hot' in a significant percentage of cases with inlet temperatures of 42° C. and above. How much of this is due to the artifacts previously noted cannot be precisely determined. A direct check is possible only in the case of 48° C. stimulation. Here an original report of warm was changed to hot or burning hot in 84.7 percent of the cases.

[8] With 48° C. without shock, hot or burning hot was reported 85.8 percent of the time; with 50° C. 96.7 percent.

Subtracting 10.2 percent for the effect of shock alone,[9] and 18.7 percent for the natural increase,[10] there still remains a percentage of 55.8 percent which appears to be directly the

TABLE 3

PERCENTAGES OF CLEAR CHANCES

(*A*) *Warm → Hot*

With	34°	40°	42°	44°	46°	48°
34°.........	5.6	9.2	18.9	23.2	31.0	46.2
25°.........	—	10.8	13.3	18.2	18.7	—
20°.........	—	2.9	8.0	5.0	16.7	—
15°.........	—	1.9	3.9	8.1	4.3	—
10°.........	—	3.5	5.3	12.5	6.1	—

(*B*) *Warm → Burning Hot*

With	34°	40°	42°	44°	46°	48°
34°.........	4.6	9.2	13.5	18.9	12.7	38.5
25°.........	—	9.7	7.2	7.6	6.3	—
20°.........	—	2.0	4.0	1.7	0.0	—
15°.........	—	1.9	2.C	5.4	8.7	—
10°.........	—	3.4	0.0	0.0	6.1	—

(*C*) *Warm → Hot or Burning Hot*

With	34°	40°	42°	44°	46°	48°
34°.........	10.2	18.4	32.4	42.1	43.7	84.7
25°.........	—	20.5	20.5	25.8	25.0	—
20°.........	—	4.9	12.0	6.7	16.7	—
15°.........	—	3.7	5.9	13.5	13.0	—
10°.........	—	6.9	5.3	12.5	12.2	—

(*D*) *Hot → Burning Hot*

With	34°	40°	42°	44°	46°	48°
34°.........	—	57.2	40.0	42.5	42.5	79.1
25°.........	—	0.0	0.0	6.7	23.1	—
20°.........	—	0.0	0.0	0.0	0.0	—
15°.........	—	0.0	0.0	0.0	8.3	—
10°.........	—	0.0	0.0	0.0	14.3	—

result of the synthesis of warm + pain. As might be expected, burning hot can also be derived from hot + pain.

2. Throughout all the tables, the addition of cold to the complex tends to mask rather than to accentuate the synthetic

[9] See 34° C. + 34° C. + shock, in Table 3-C.

[10] See Table 2.

effects. Except with 25° C., the results are negligible. The only known artifact here is the effect of the cold bars upon the surface temperature of the warm bars. In Table 1, it was shown that the reports of 'hot' with simple stimulation are not significantly fewer when the cold bars are at 25° or 20° C. instead of 34° C. Yet clearly the compound effects are much lower with 25° C. and almost disappear at 20° C.

TABLE 4

Subjects Reporting Hot and Burning Hot

	With Simple Stimulation plus Shock	With Compound Stimulation plus Shock
Warm → hot	65	28
Warm → burning hot	44	23
Warm → hot or burning hot	74	37
More than 50 percent of clear chances	35	8
Hot → burning hot	49	6

Table 4 presents the situation from another angle. 74 out of 108 subjects reported the change from warm to hot or to burning hot at least once with simple stimulation plus shock. 35 of these gave such reports in more than half of the clear chances; 8 of them in 100 percent of the cases. In contrast, only 37 subjects reported changes from warm to hot or burning hot with compound stimulation plus shock, and only 8 of them in more than half of the clear chances. Significantly, 7 of these 8 subjects gave equal or better records with simple stimulation plus shock.

Discussion of Results

With untrained subjects, it is now evident that reports of 'hot' and 'burning hot' can be secured from a variety of stimulus situations.

Reports of 'hot' can be obtained:

(*a*) Frequently with warm + shock.
(*b*) Occasionally from warm + cold.
(*c*) Rarely from neutral + shock.

Reports of 'burning hot' can be obtained:

(*a*) Frequently with hot + shock.

(*b*) Less frequently with warm + shock.
(*c*) Occasionally with warm + cold + shock.
(*d*) Rarely with neutral + shock.

This gives us a new light on the synthetic approach with untrained subjects. If 'hot' and 'burning hot' can be obtained from so many stimulus combinations, what is the value of any *one* of them. What point would there be, for example, in attempting to develop a new technique of compounding warm and cold which might give positive results? Evidently cold is not a *necessary* condition for the synthetic experience of 'hot,' because it can be obtained in abundance without cold.

Does this prove that the Alrutz theory (as applied to the common experience of 'hot') is wrong? Not at all! The negative results with warm + cold, coupled with the positive results from warm + shock, may well make us view the theory with some suspicion. Clearly, the synthetic 'hot' occurs more readily in the absence of cold . . . but nevertheless cold receptor stimulation may be a necessary part of the *normal* experience of 'hot.' Only physiological analysis can give us the answer.

The positive results with warm + shock make us curious as to what might happen with trained observers under similar conditions. Can the trained observers' 'heat' be synthesized from the warm + shock, as well as from warm + cold? This is a question which only further research can answer.

Summary

108 untrained subjects were tested with a variety of compound temperature stimulations plus mild electric shock. With simple warm stimulation plus shock, clear reports of 'hot' and 'burning hot' were obtained. When cold was added to the complex, the positive effects fell off sharply and were virtually negligible at the lower temperatures. It has now been shown that 'hot' can be synthesized quite readily from warm + shock, infrequently from warm + cold, and occasionally from neutral + shock. It is therefore main-

tained that the synthetic approach can give us no light upon the normal physiological basis for the common experience 'hot.' Because of the largely negative results with warm + cold, and the positive results with warm + shock without cold, it is suspected that the Alrutz theory (simultaneous stimulation of warm and cold receptors) does not apply to the common experience of 'hot.' But the final test must be analysis, not synthesis. It is suggested that similar experiments with introspectively-trained observers would be desirable.

(Manuscript received December 31, 1937)

Reprinted from the JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
Vol. 23, No. 4, October, 1938
Printed in U. S. A.

STUDIES IN THERMAL SENSITIVITY. 7. FURTHER SYNTHETIC EVIDENCE AGAINST THE ALRUTZ THEORY

BY WILLIAM LEROY JENKINS

Lehigh University

Two previous studies in this series have presented the results of synthetic experiments which indicate that the Alrutz theory of 'heat' (simultaneous warm plus paradoxical cold) does not apply to the common experience 'hot.' Warm + cold grill stimulation of naive subjects does not result in consistent reports of 'hot,' which actually occur *less* frequently with this compound than with the corresponding simple warm stimulation alone.[1] 'Hot' and 'burning hot' can be synthesized from simple warm + electric shock in a fair number of instances, but the addition of cold to the complex sharply reduces the positive effects.[2]

Following an informal presentation of these results, objections were raised to the *technique* on two grounds: (1) Flat bar grills had been used in these studies, whereas the original synthetic experiments with trained observers had been performed with grills made of round tubing.[3] (2) Subjects had not been instructed to press down hard on the grill, and strong pressure was said to aid the synthesis. The best answer to these objections is the outcome of direct experimental tests.

[1] Jenkins, W. L., Studies in thermal sensitivity: 5. The reactions of untrained subjects to simultaneous warm+cold stimulation, *J. Exper. Psychol.*, 1938, **22**, 451–461.

[2] Jenkins, W. L., Studies in thermal sensitivity: 6. The reactions of untrained subjects to simultaneous warm + cold + electric shock, *J. Exper. Psychol.*, 1938, **22**. 564–572.

[3] For economy of space, these will hereafter be referred to as the 'Dallenbach experiments,' since the best of them were carried out in the Cornell laboratory under the guidance of Dr. K. M. Dallenbach. For a bibliography and brief review, see: Jenkins, W. L., Studies in thermal sensitivity: No. 5. The reactions of untrained subjects to simultaneous warm + cold stimulation, *J. Exper. Psychol.*, 1938, **22**, 451–453.

Apparatus and Procedure

Two additional grills were constructed for this work: No. 5, of 3 mm diam. copper tubing with 1½ mm air gap between tubes; No. 6 of 4½ mm diam. and 2 mm air gap. In both grills, the total stimulating surface was approximately 8 × 7 cm, and the cold tubes could be raised and lowered, permitting a direct comparison between simple warm stimulation and the same warm compounded with cold. Grill No. 1, with 3 mm flat bars, was used as a control.

One pair of temperatures (40° C. and 20° C.) was employed throughout. Each subject was tested 10 times with each of the three grills, with the simple stimulation preceding the compound; and also 4 times with Grill No. 1 when the compound stimulation was given first. Half of the stimulations were made with light pressure and half with heavy pressure, in random order. For light pressure, the subject merely allowed the volar surface of his left forearm to rest upon the grill. For heavy pressure, he was instructed to bear down as hard as he could. All grills were concealed by a hood. Reports were made in terms of: cold, neutral, warm, hot or burning hot.

Results

Table 1 shows that the results were as flatly negative as

TABLE 1
Percentages—Reports of 'Hot'

Grill No.	Pressure	Simple 40° Alone	Compound—40° and 20°	
			Total	Clear
1	Light	15.5	0.0	0.0
1	Heavy	16.2	0.7	0.9
1*	Light	6.9	10.3	12.3
1*	Heavy	5.2	5.2	5.5
5	Light	10.7	2.1	2.4
5	Heavy	12.8	2.1	0.8
6	Light	9.3	3.1	3.4
6	Heavy	12.4	2.5	2.8

* Compound stimulation given first.

those obtained in the earlier study, neither tubular grills nor heavy pressure having any significant effect. Heavy pressure increased the reports of 'hot' from simple stimulation, but actually reduced the small number obtained with the compound. When the compound stimulation was given first, the 'clear hot' reports with light pressure rose to 12.3 percent. But even this figure must be viewed with suspicion, because the simple stimulation itself gave 15.5 percent.

Table 2 confirms the analysis. 19 of the subjects never gave a single clear report of 'hot' with compound stimulation.

Table 2

Clear Reports of 'Hot' by Subjects

Never	19
0–9 percent	8*
10–24 percent	1*
25–29 percent	1
Over 30 percent	0
Total	29

* In 7 of these 9 cases, 'hot' was reported a greater number of times with simple stimulation.

Nine others did report it sporadically, but 7 of these called the simple stimulation 'hot' more frequently. There is just one apparently genuine case of synthesis . . . the individual who reported 'hot' with 26 percent of the compound stimulations and never called the simple stimulation 'hot.' A few similar instances were found in the earlier study.

Discussion

These experimental tests seem to dispose of the objections on the grounds of technique. Considering that negative results were also secured by Dallenbach in the single instance when a naive individual was tested,[4] it appears unlikely that further variations of warm + cold synthesis would yield any significant number of reports of 'hot' from untrained subjects.

Another objection which has been raised to these negative results is that they are *meaningless*. Naive individuals, it is argued, do not report 'heat' in response to simultaneous warm + cold stimulation, because they have not been properly trained to recognize 'heat' when they encounter it. Hence, their testimony on the subject is worthless.

This objection is cogent only if we assume that there is one universal experience from high temperature stimulation. But why should we make this assumption? On both physiological and psychological grounds, there is reason to expect a variety of experiences. Physiologically, the warm, cold and pain receptors might respond differently in various individ-

[4] Ferrall, S. C. and Dallenbach, K. M., The analysis and synthesis of burning heat, *Amer. J. Psychol.*, 1930, 42, 76.

uals. Psychologically, in the perception called 'heat' or 'hot,' the roles played by warm, cold and pain excitation might vary in different people—even if the physiological substrate was identical in all cases.

When we approach the literature from this viewpoint, we can distinguish at least *five* variant definitions of 'heat' or 'hot.' These may be briefly summarized as follows:

1. For Goldscheider, 'heat' was merely a higher degree of 'warm,' but because of the long-lasting after-sensations, he held that stimulation of the *Gefühlsnerven* was also involved. Cold sensations might or might not be present, but these were adventitious and could be analyzed out of the mixture.[5]

2. For Kiesow, the participation of the pain receptors was an essential part, of what he called 'heat.' where the experience was not actually painful, it was always *schmerzbetont.* Cold receptor excitation was not essential, because 'heat' could be experienced in regions where no cold receptors could be demonstrated.[6]

3. For Alrutz, 'heat' was a unique and unanalyzable experience, easily distinguishable from 'warm.' Its critical feature, he held, was the participation of cold receptor excitation, because 'heat' could not be experienced where cold receptors were absent.[7]

4. For Hacker and also for Thunberg, although cold receptor excitation was involved, 'heat' was *not* a unique and unanalyzable experience. Rather, it was a *Mischempfindung* in which the warm and cold elements could be distinguished with practice.[8]

5. In addition to these four different 'heat' experiences

[5] Goldscheider, A., Beiträge zur Lehre von der Hautsensibilität, *Zsch. f. klin. Med.*, 1912, **75**, 9–10.

[6] Kiesow, F., Zur Analyse der Temperaturempfindung, *Zsch. f. Psychol. u. Physiol. d. Sinnesorg.*, 1901, **26**, 238.

[7] Alrutz, S., Wie man die Natur der Hitzempfindung beweist und demonstriert, *Ber. ü. d. 6 Kongr. f. exper. Psychol.*, 1914, 16,

[8] Hacker, F., Beobachtungen an einer Hautstelle mit dissozierter Empfindungslähmung, *Zsch. f. Biol.*, 1913, **61**, 242. Thunberg, T., Untersuchungen über die relative Tiefenlage der kälte-, wärme- und schmerzperzipierenden Nervenden in der Haut und über das Verhältnis der Kaltenervenden gegenüber Wärmereizen, *Skand. Arch. f. Physiol.*, 1901, **11**, 415.

described by trained research workers, it seems reasonable to grant the existence of a fifth, which we have termed the common experience 'hot.' For most naive individuals, 'hot' is merely a *pseudo-discrete* step in an actual continuum. 'Barely warm' changes to 'warm' and to 'very warm' and to 'hot,' without any sharp dividing lines. The numbers 1, 2, 3, 4, etc. would serve equally well. As far as the present evidence shows, cold excitation plays no essential role.

From this varied array of experiences, we may choose that of Alrutz and arbitrarily label it 'true heat.' Then, quite obviously, only those who have this particular experience are qualified to take part in experiments concerning its nature. This eliminates the evidence of Goldscheider, Kiesow, Hacker and Thunberg, because these men did not experience 'true heat.' The same objection applies to studies with untrained subjects.

Fortunately, the trained observers in the Dallenbach experiments did experience 'true heat'—unique and unanalyzable, qualitatively different from 'warm'—and this could be synthesized consistently from simultaneous warm + cold stimulation. Unfortunately, such synthetic evidence does not prove that cold excitation is a necessary part of the natural experience. Somewhat disturbing are the casually-reported findings of Knight, who synthesized 'burning heat,' 'burn' and 'hotness' from punctiform warm + pain in the same laboratory. Is it possible that the 'synthetic heat' from warm + cold is merely an illusion? Final judgment must await the results of analytic tests to show whether 'true heat' can be experienced where no cold sensitivity exists.

Meanwhile, there remains the problem of what to do about the common experience 'hot.' Although it is clearly not 'true heat,' it seems to be the experience of the bulk of ordinary individuals. Hence, its constitution may be of interest to general psychology. Does the Alrutz theory of warm plus paradoxical cold excitation apply? The negative results with attempted warm + cold synthesis provide a strong presumption (but not proof) that it does not. The final answer must come from critical analytic tests.

Summary

Additional experiments with simultaneous warm + cold stimulation, employing tubular grills with light and heavy pressure, confirm the negative results previously secured with flat bar grills. Untrained subjects do not report 'hot' in response to warm + cold stimulation. To the objection that these results are meaningless because naive individuals have not been properly trained to recognize 'heat' . . . the following answer is made: Five clearly-different experiences called 'heat' or 'hot' have been described in the literature, four of them by trained investigators. We may arbitrarily select the experience of Alrutz and label this 'true heat.' But the common experience 'hot' is also of interest to general psychology. According to the synthetic findings, cold excitation is probably not an essential element and the Alrutz theory does not apply. Final judgment must wait upon more critical analytic evidence.

(Manuscript received April 23, 1938)

Reprinted from the JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
Vol. 23, No. 4, October, 1938
Printed in U. S. A.

STUDIES IN THERMAL SENSITIVITY. 8. ANALYTIC EVIDENCE AGAINST THE ALRUTZ THEORY

BY WILLIAM LEROY JENKINS
Lehigh University

Three preceding articles in this series have presented the results of synthetic experiments which indicate that the Alrutz theory of 'heat' (simultaneous warm plus paradoxical cold) does not apply to the common experience 'hot.' With untrained subjects, 'hot' cannot be synthesized from simultaneous warm + cold stimulation with flat bar grills,[1] nor with tubular grills regardless of whether light contact or heavy pressure is employed.[2] 'Hot' and 'burning hot' can be derived in a fair percentage of cases from simple warm + electric shock, but the addition of cold tends to destroy the effect.[3]

Neither these negative results with untrained subjects nor the positive results obtained by Dallenbach and others provide critical evidence concerning the constitution of 'hot' or 'heat.' The negative results may be an artifact of faulty technique. The positive results may be illusory. Whether cold excitation is a necessary part of the *natural* experience can be checked only by direct analytical attack.

Analytically, the critical question is: "Can 'hot' (or 'heat') be experienced in skin areas where cold sensitivity is either absent or not functioning?" If 'hot' (or 'heat') can be consistently experienced under these conditions, then cold excitation is clearly not a *necessary* part of the experience and the Alrutz theory does not hold.

[1] Jenkins, W. L., Studies in thermal sensitivity: 5. The reactions of untrained subjects to simultaneous warm + cold stimulation, *J. Exper. Psychol.*, 1938, **22**, 451–461.

[2] Jenkins, W. L., Studies in thermal sensitivity: 7. Further synthetic evidence against the Alrutz theory, *J. Exper. Psychol.*, 1938, **22**, 411–416.

[3] Jenkins, W. L., Studies in thermal sensitivity: 6. The reactions of untrained subjects to simultaneous warm + cold + electric shock, *J. Exper. Psychol.*, 1938, **22**, 564–572.

No accounts of systematic analytic studies of this nature are found in the literature. Alrutz states categorically that 'heat' cannot be experienced in regions where cold sensitivity is absent, with which Hacker and Thunberg agree, although their definition of 'heat' is not the same as his. Goldscheider and Kiesow are equally positive that 'heat,' as they use the term, can be experienced in regions where no cold sensitivity can be demonstrated. In no case, however, is detailed evidence reported. Nor were such systematic tests carried out with the introspectively-trained observers of the Dallenbach experiments. The work to be reported here is an attempt to answer the critical question only for the common experience 'hot.'

Procedures

1. *With Cold-adapted Areas.*—Subjects were divided into groups of three, each individual acting in turn as experimenter, subject and recorder. Three circles (13 mm diam.) were stamped on the volar surface of the left forearm. A stimulator 13 mm in diam., maintained at 18–20° C. by circulating water, was held against the arm within a marked circle for three minutes (unless complete adaptation occurred sooner). A stimulator 2 mm diam. at 50° C. was then applied at random inside and outside of the circles for a total of 10 stimulations each, the subject having his eyes closed. This was repeated in rotation, using a different circle each time, until each subject had received 60 stimulations (30 inside and 30 outside of the circles). The subjects reported: cold, neutral, warm, hot or burning hot.

2. *With Cold-mapped Areas.*—This experiment was also carried out in groups of three. The usual precautions in mapping were observed, and the members of each group rotated positions every few minutes to minimize the effects of adaptation. In an area 2 × 2 cm, the parts sensitive to cold (2 mm diam. stimulator at 18–20° C.) were marked directly on the skin. The stimulator was then brought to 50 C. and each mapped area was stimulated in random order for a total of 10 stimulations within the cold-sensitive parts and 10 in the non-cold-sensitive. This was repeated in rotation for a total of 60 stimulations for each individual (30 in cold, and 30 in non-cold). The subjects reported: cold, neutral, warm, hot and burning hot.

3. *With Seriatim Stimulation.*—In this procedure, the author acted as experimenter and the subjects were tested in groups of 3 to 8. A block of 25 spots (each about 3 mm square and about 3 mm apart) was stamped on the volar surface of the left forearm. The 25 spots on each subject were stimulated in random order with a 3 mm diam. stimulator, the members of the group being handled in rotation to minimize the effects of adaptation. Low temperature stimulation was alternated with high temperature stimulation, until 6 low (3 at 20° C. and 3 at 10° C.) and 5 high (the same throughout, but varying between 40° C. and 50° C. with different groups) had been completed upon each subject. In one set (49–50° C.), only 5 cold stimulations, all at 20° C., were used.

Results

1. In the first procedure, the idea was to render the cold receptors inoperative through long-time adaptation. 56 subjects were stimulated with 50° C. 1,614 times within the cold-adapted circles and an equal number of times on the normal skin outside. 'Hot' or 'burning hot' was reported 355 times from within the circles and 462 times from outside. Three objections might be raised to these results: (1) There is no direct evidence that cold-adaptation was complete as far as paradoxical stimulation is concerned. (2) Just such results might be expected if the subjects were reporting 'warm' and 'hot' at random without regard to the actual experience. (3) On the other hand, the positive results might be reduced an undetermined amount by the effect of the prolonged pre-cooling on the warm receptors. In fact, this is indicated by the reduced *total* number of warm and hot reports . . . 1,049 from outside and 930 from inside the circles.

2. In the second procedure, the idea was to compare the results from cold and non-cold-sensitive areas. 73 subjects were stimulated with 50° C. 2,190 times in non-cold-sensitive parts and an equal number in the cold-sensitive parts. 'Hot' or 'burning hot' was reported 560 times from the former and 618 times from the latter. Three objections might be raised to these findings: (1) With untrained subjects, the cold-mapping was undoubtedly inaccurate. (2) Figures like these could result from purely random reporting. (3) There might be cold receptors in the area which responded to paradoxical stimulation, but not to normal stimulation at 18–20° C.

3. The third procedure was designed to meet these objections. The seriatim scheme of stimulation provides evidence of the consistency of reports from any given spot, eliminating the possibility that the results could be due to random reporting. Furthermore, if there are actually some cold receptors which respond better to paradoxical than to normal stimulation, a considerable number of reports of paradoxical cold should be obtained from spots that were neutral to normal cold stimulation.

Table 1 shows the gross results for 64 subjects according to three different criteria of consistency. Spots that were consistently 'hot' (to high temperature stimulation) were relatively as frequently 'neutral' (to low temperature stimulation) as they were 'cold' (to low temperature stimulation). With perfect consistency as the criterion, the percentages are almost identical. With less exacting criteria, the equality relationship is unchanged.

Concerning the possibility of spots which respond to paradoxical but not to normal low temperature stimulation, the record is equally clear. Out of 10,625 stimulations at temperatures between 40° C. and 50° C., paradoxical cold was reported exactly 16 times in all—in every case from spots which were consistently 'cold' to normal stimulation. There is not even a single instance of a report of paradoxical cold from a spot which was consistently neutral to normal low temperature stimulation.[4]

TABLE 1

SPOTS CONSISTENTLY CALLED 'HOT' AS COUPLED WITH 'COLD' AND WITH 'NEUTRAL'

Criterion of Consistency		'Hot' Coupled with 'Neutral'	'Hot' Coupled with 'Cold'
For 'Hot'	For 'Cold' or 'Neutral'		
5/5	6/6	44/140 *31.4%*	145/482 *30.2%*
4/5	5/6	121/288 *42.0%*	340/787 *43.3%*
3/5	4/6	203/438 *46.3%*	499/1008 *49.6%*

Note: Percentages are computed in terms of total spots in each group giving positive responses to high temperature stimulation. For example: 140 spots were positive 5 times out of 5 high temperature stimulations and 'neutral' 6 times out of 6 low temperature stimulations. Of these 140 spots, 44 were 'hot' 5 times out of 5 high temperature stimulations—31.4 percent. The other figures in the table were similarly derived.

[4] Reports of paradoxical warm occurred 85 times out of 4,650 stimulations at 20° C., and 25 times out of an equal number at 10° C. In spite of the comparative rarity of both paradoxical phenomena, there is some evidence of their genuineness. On two occasions, paradoxical cold was reported 3 times from the same spot. In two cases, paradoxical warm was reported 4 times from the same spot; and in 8 cases, 3 times from the same spot.

Discussion

Unless serious exception can be taken to all three of the procedures reported, it seems quite clear that cold excitation is not necessary for the common experience 'hot.' Untrained subjects report 'hot' with approximately equal frequency from 'cold' spots and from 'neutral' spots. Thus the provisional conclusion from the negative results with warm + cold grill stimulation is confirmed. *The Alrutz theory does not apply to the common experience 'hot.'*

It will be readily admitted that these subjects were not experiencing 'heat' . . . in the Alrutz sense of the term. In fact, that is exactly our point. The Alrutz 'heat' is not the same as the common 'hot.' As Alrutz described it, 'heat' is unique—qualitatively different from 'warm.' As the untrained subject describes it, 'hot' is merely a higher degree of 'warm' and not qualitatively different. This latter appears to be the experience of most naive individuals.

Further research may reveal that some individuals *spontaneously* have the experience of Alrutz from high temperature stimulation. Again, it may be possible to *teach* individuals to have such a perception, although they did not have it originally. In either event, the acid test is not the effect of warm + cold synthesis, but rather some analytical technique, preferably seriatim stimulation of a large number of marked spots. Then, if cold excitation is essential for the experience, the subject should never report 'heat' from a spot that is consistently neutral to low temperature stimulation.

It seems fair to challenge the supporters of the Alrutz theory to perform such experiments with introspectively-trained observers. When a trained observer states that 'heat' is unique and qualitatively different from warm, that is a subjective judgment. When a trained observer says that his experience from simultaneous warm + cold grill stimulation is identical with 'heat,' that is another subjective judgment.

But when a trained observer holds that 'heat' is *dependent* upon paradoxical cold excitation, that is clearly a matter for objective check. He can submit himself to the crucial test of seriatim stimulation. If he reports 'heat' from any spots

that are consistently neutral to cold stimulation, then cold is not a necessary part of his experience—and the Alrutz theory does not apply to it. If he reports 'heat' only from spots which are consistently cold to low temperature stimulation this does not necessarily prove the Alrutz theory. It remains to be shown that paradoxical stimulation of cold receptors does actually occur every time 'heat' is experienced. But this is merely tilting at windmills. Let us first see whether trained observers, working without knowledge of which spots are cold and which are neutral, will report 'heat' only from 'cold' spots and never from 'neutral' spots. That is the next step to be taken.

Meanwhile, it is safe to say that the Alrutz theory should not be stated as a generalization. It clearly does not hold for the common experience 'hot.' From a critical standpoint, there is absolutely no direct, objective evidence that it has ever applied to anyone but Alrutz himself.

Summary

Three procedures were used to check whether cold excitation is necessary for the common experience 'hot': (1) reports of 'hot' from cold-adapted and normal skin; (2) reports of 'hot' from mapped cold and non-cold-sensitive areas; (3) seriatim stimulation of 25 marked spots, alternately with high and low temperatures. Results from all three procedures were in agreement: 'hot' is reported whether 'cold' is present or not. Clearly, cold excitation is not a *necessary* part of the common experience 'hot' and the Alrutz theory does not apply to it. Supporters of the Alrutz theory are challenged to apply similar tests to trained observers. There is no objective evidence which shows that the Alrutz theory applies even to what trained observers call 'heat.'

(Manuscript received April 23, 1938)

LEHIGH UNIVERSITY STUDIES

Psychology Series, No. 3

Evidence for a "Concentration Theory" of Cutaneous Cold Sensitivity

by

WILLIAM LEROY JENKINS
Assistant Professor of Psychology

Reprinted from the Journal of Experimental Psychology

LEHIGH UNIVERSITY
BETHLEHEM, PENNSYLVANIA
1939

Lehigh University Studies comprise selected papers arising from the creative work of the University which for any reason are not issued as periodical bulletins from the **Institute of Research.**

LEHIGH UNIVERSITY STUDIES

Science and Technology Series

No. Title

* 1. Tables of the Properties of Aqua-Ammonia Solutions. 15 cents.

Psychology Series

No. Title

* 1. Learning to Generalize and An Experiment in Generalizing a University Problem. 15 cents.

* 2. Evidence Against the Alrutz Theory of Heat. 15 cents.

* 3. Evidence for a "Concentration Theory" of Cutaneous Cold Sensitivity. 15 cents.

* Limited number of copies available for free distribution.

Reprinted from the JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
Vol. 24, No. 3, March, 1939
Printed in U. S. A.

STUDIES IN THERMAL SENSITIVITY: 9. THE RELIABILITY OF SERIATIM COLD-MAPPING WITH UNTRAINED SUBJECTS

BY WILLIAM LEROY JENKINS
Lehigh University

STUDIES IN THERMAL SENSITIVITY: 9. THE RELIABILITY OF SERIATIM COLD-MAPPING WITH UNTRAINED SUBJECTS

BY WILLIAM LEROY JENKINS

Lehigh University

A knowledge of the consistency with which thermal sensitivity can be mapped is of interest, not only for itself, but also as a basis for further research. If reasonably high reliability can be obtained with repeated mappings, it indicates: (*a*) that the sensitivity of the skin is relatively stable; (*b*) that the experiences can be consistently reported.

The importance of the problem was recognized some years ago by Dallenbach.[1] Impressed by the poor results obtained with repeated mappings in their student laboratory, he set out to discover and control all of the significant variables. A long series of preliminary tests disclosed a dozen of these which previously had been largely neglected. The final procedure which he developed stands as a model of technical excellence. The three observers were given a long course of training before the start of the main experiments, in which a stamped grid was carefully maintained on the upper arm of each observer and mapped for cold and warm at the same hour of each of the four days (Monday, Wednesday, Friday and Monday).

The maps for cold (1 mm diam. stimulator at 8–9° C.) are shown in the upper half of Fig. 1. Even casual inspection reveals a certain family resemblance among the four maps for a given area, but a statistical statement of reliability is harder to achieve. High degrees of correspondence are indicated by the 'student method' and by the 'absolute

[1] Dallenbach, K. M., The temperature spots and end-organs, *Amer. J. Psychol.*, 1927, 39, 402–427.

method' of computation, but Dallenbach himself points out that neither method is satisfactory.[2]

A graphic representation shows better the true nature of the findings. At the bottom of Fig. 1 are large composite

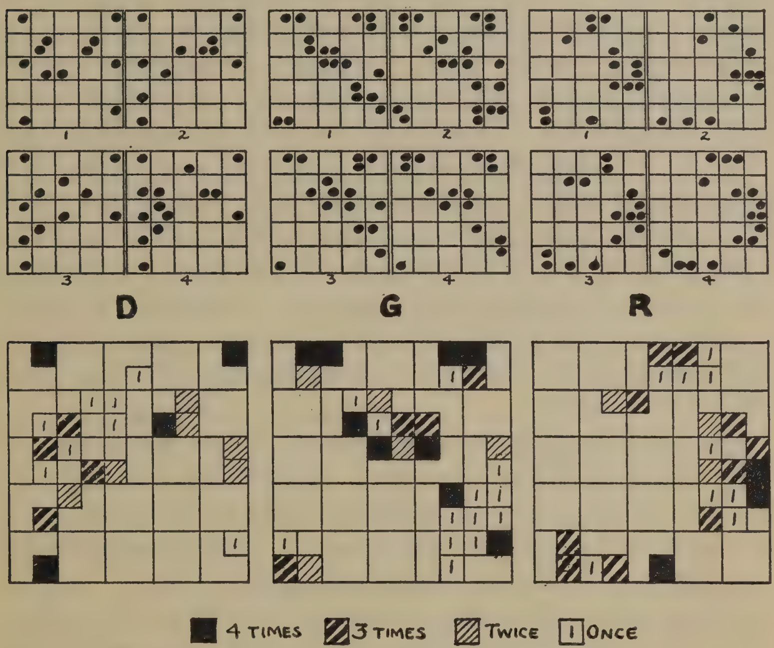


FIG. 1. Replotting of Dallenbach's cold-mapping.

maps on which are indicated the *number of times* each mm square was reported 'cold' during the four trials. Among the 22 positive squares on D's map, for example, 4 were re-

[2] Statistically, both methods are practically worthless, because high scores might be obtained by pure chance. In the 'student method,' the correspondence of one map with another is determined by the percentage of spots which are identically localized or fall within 1 mm of one another. This actually means that each square on the first map 'corresponds' to 9 squares on the second. Thus, it would be possible to scatter 15 spots on the first map so that they completely blanketed all the squares on the second. Then, no matter how random the arrangement on the second map, a 'correspondence' of 100 percent would be forced. In the 'absolute method,' the identical squares on the two maps are counted and this divided by 100 gives the correspondence. An example will show the fallacy. Imagine 5 spots down the left-hand edge of one map and 5 down the right-hand edge of the other. Obviously, the true correspondence is nil, but the 90 identical *blank* squares would give a fictitious correspondence of 90 percent.

ported cold on all four occasions, 4 three times, 6 twice and 8 only once. Or, considering the total of 48 reports of cold during the four days, 28 could be called 'consistent' (reported 3 or 4 times from the same square); whereas the remaining 20 were reported only once or twice from the same square. Similar results were obtained with observers G and R.

Such a large percentage of 1's and 2's provides no very satisfactory evidence of consistency . . . on the face of it, quite the contrary. We are faced with a dilemma. If 'cold spots' are stable entities, the mapping must be at fault. The experimenter must have misplaced his stimulator rather frequently, or the observer must have made numerous errors in reporting. If the mapping is satisfactory, the so-called 'cold spots' must be highly inconstant. Even if we allow for an occasional error on the part of the experimenter or the observer, the great majority of them appear to be here today and gone tomorrow.

Are such inconsistent 'cold spots' worth considering as entities at all? What is the point of declaring that there are (say) 12 cold spots to a square cm, if the 12 we locate today are not the same 12 we found yesterday . . . or if an average of 12 means a total of 22 in four mappings?

Let us discard for the moment the concept of so many discrete cold spots per square cm. Let us examine the composite maps simply as charts which indicate roughly the parts of the area which respond to a 1 mm stimulator at 8–9° C. The first thing we notice is that there are no *isolated* squares anywhere on the maps, except at the very edges where we cannot know what lies beyond. Then we observe that the marked squares are grouped into sizable clusters. Finally, we might speculate that the 4's and 3's in these clusters look remarkably like peaks of sensitivity, with the 2's and 1's indicating less responsive regions.

Some of the apparent anomalies of the maps now make perfectly good sense. For instance, the patch of nine 1's in the lower right-hand corner of G's map no longer needs to be considered as a concentrated case of poor work on the part of experimenter and observer. It becomes a 'valley' between

twin peaks. Likewise, the string of seven 1's in the upper right-hand part of D's map fits in perfectly well with the general conformation of that region.

Viewed in this way, all three of the maps show an orderly, if non-uniform, scheme of distribution . . . which implies consistency in the mapping. Unfortunately, there is no direct proof. We have explained away the apparent unreliability of the mapping by assuming a certain non-uniform scheme of distribution in the skin sensitivity itself. We have no final evidence to show that this hypothesis is correct.

The ambiguous outcome of Dallenbach's study points to the need for a different technique. A single mapping at each session is not sufficient. *Seriatim* stimulation during each period is required to establish independently the reliability of the mapping. Every square must be tested, not once, but a number of times at each session. Then comparison on a statistical basis will be possible.

Dallenbach's research was performed upon three highly-trained observers. What sort of results might we expect from totally-untrained subjects? Will their performance really be as poor as the usual student laboratory work indicates? Is it possible that, under fair conditions, completely untrained individuals might show good records of consistency?

These are not purely academic questions. In many instances, it is possible to secure a large number of student-subjects for a few hours apiece, but difficult to arrange for long training programs. If initially-good subjects can be found in such a group, valuable data might be obtained with only the limited training available. Or again, large numbers of individuals might be studied in a minimum of research time—for example, to outline the range of individual differences in some phase of thermal sensitivity. In any event, untrained subjects should show us the very worst we can expect from the seriatim method as a research technique. If it yields satisfactory results with them, there can be no question of its value when applied to those with ample training.

Procedure

The chief features of the procedure may be summarized as follows:

1. *Skin-marking.*—Instead of the traditional grid, a checkerboard of 50 squares was stamped on the skin with printer's ink. This was found to aid accurate placement of the stimulator, and had the additional advantage that it could be renewed if necessary. Although the whole figure might be distorted because of the elasticity of the skin, the individual squares could be restamped one by one with a minimum of displacement.

2. *Stimulator Sizes.*—Stimulators with tip diameters of 2 mm and 3 mm were used. Preliminary experiments showed that the smaller sizes rarely gave a sufficient range of discriminable experiences. The 1 mm size used in the Dallenbach study, for instance, results in a very large percentage of 0's in most subjects, which is undesirable in a study of reliability. If we discard provisionally the concept of 1 mm 'cold spots' as sacred entities, there is no particular point in using the 1 mm size.

3. *Seriatim Method.*—During each one-hour period, every square in the checkerboard pattern was stimulated 6 times, at intervals of 7–10 minutes. An ordinary brass hand-stimulator was used, with no particular device to insure uniform pressure. The temperature (17° C.) was maintained within 1° C. by rapidly circulating water.

4. *Reports.*—Subjects were instructed to report in three categories: 2 for strong cold, 1 for weak cold and 0 for neutral. This scheme was used in place of simply 'cold' and 'neutral' to provide a broader base for scoring. Subsequent events indicated that a still larger number of categories might have been used.

5. *Order of Stimulation.*—To avoid any possibility that the subject might memorize a sequence of reports, an approximately random order of stimulation was employed. The black squares and the white squares of the checkerboard pattern were stimulated separately, in rows of five at a time. The sequence of these rows was changed with each trial, and sometimes the stimulation proceeded from right to left in a row, sometimes the reverse.

6. *Individual and Group Experiments.*—In the individual experiments, the author acted as experimenter and two students alternated as subject and recorder, for two consecutive one-hour periods, making 12 rounds in all. In the group experiments, a student acted as experimenter for each group, while two others alternated as subject and recorder. The group schedule was:

> 1A and 1B—Consecutive periods Monday afternoon.
> 2—One hour Tuesday morning.
> 3—One hour Wednesday morning.
> 4A and 4B—Consecutive periods Thursday morning.

Only two of the 15 subjects (4 women, 11 men students) were able to complete all six periods, but the majority came for four periods during which the same checkerboard pattern was maintained (renewed if necessary) and stimulated six times at each period. In both the individual and the group experiments, the marked area was on the volar surface of the left forearm.

7. *Statistical Treatment.*—The three categories 2, 1 and 0 are ordinal, not cardinal numbers. However, the best statistical assumption seems to be that '1' lies half-way between '0' and '2.' Then a quantitative score for each square can be obtained simply

by adding the numbers reported from that square during a period. The correlation between any two periods can be computed by the product moment formula.[3]

Variables and Artifacts

The reliability coefficients thus obtained will necessarily be *resultants* of a number of possible variables which seem to be inherent in the situation. These include: (*a*) changes in the sensitivity of the area as a whole; (*b*) changes in the responsiveness of individual squares; (*c*) general ability of the subject to maintain standards of judgment during a period; (*d*) shifts in the level of criteria from period to period; (*e*) improvement of discrimination with practice. Unfortunately, it is not possible to isolate the individual influences of these variables by any gross consideration of results. For example, a change in the sensitivity of the area as a whole has exactly the same effect on the scores as a shift in the general level of the subject's criteria. Again, changes in the sensitivity of individual squares produce irregularities which are duplicated by the subject who cannot maintain consistent standards of judgment.

In some cases, the detailed evidence is quite suggestive. For instance, a square might show 6 reports of '2' at one period and 6 reports of '1' at another, with no parallel change in the general level of scores. Clearly, each score alone is perfectly consistent; so the chances are that the sensitivity of the square itself has changed. On the other hand, when a number of squares show mixtures of 2's, 1's and 0's, it is probable that we are dealing with a poor subject. Sharpening of discrimination with practice is indicated by an increasing tendency of the scores to group closely around 0, 6 and 12. At the start, many intermediate scores result because a square is sometimes called '2' and sometimes '1'; or sometimes '1' and sometimes '0.' As the subject develops more skill and the boundaries between categories become more precise, these borderline cases tend to fall consistently into one category or the other. Although this is a highly desirable development, it results in a reduction in the computed reliability coefficient when early results are compared with those late in the series.

In addition to these variables which seem to be inherent in the situation, there are a number of possible *artifacts* which can be wholly or partially controlled. Some of these would tend to increase, some to decrease the computed coefficients.

Artifacts which would tend to *increase* the coefficients include:

1. Adoption of a fixed sequence of reports by the subject, regardless of the actual results of stimulation. This was entirely eliminated, it is hoped, by the random order of stimulation.

2. Cues from the experimenter. This probably played no part, since the experimenter himself usually did not know what the previous reports from the square had been.

3. Collusion between subject and recorder. Since subject and recorder were usually in keen competition, this is highly improbable in the present instance.

---

[3] Although a surprisingly large number of cases of 'six of a kind' occurred, these are not necessary in order to establish consistency. For example, a square with reports of 2-1-1-2-1-2 in one period and 2-2-1-1-2-1 in the second is perfectly consistent. The only questionable cases are those in which the same score is arrived at by a different combination of numbers. This is possible only when the extreme reports (2 and 0) occur in the same square during the same period. In the better subjects, this was relatively rare. Typically, for the intermediate scores, 2's were mixed with 1's and 1's mixed with 0's.

4. Artificial elimination of one category by the subject. Perhaps the best answer to this is that practically all of the highest coefficients were obtained by subjects who clearly used the full range of categories.

Among the artifacts which would tend to *decrease* the obtained coefficients are:

1. The effects of repeated stimulation upon sensitivity—especially the negative effect of adaptation. It may seem strange that 2-second stimulations 7–10 minutes apart could have any such influence, but this seems clearly indicated in some cases.

2. Mis-location of the squares in restamping. This would enter only in that minority of cases where the markings faded to such an extent that they had to be renewed.

3. Errors of the experimenter, including misplacement of the stimulator, uneven pressure and failure to get the stimulator tip flat against the skin. Especially with untrained experimenters, such artifacts must have entered to some extent.

4. Errors of the recorder, such as writing down an incorrect number of making the entry in the wrong place on the data sheet. This probably occurred occasionally.

In summary, it seems probable that the sum total of the artifacts would tend to *decrease* rather than to increase the computed coefficients. It is safe to say, therefore, that the reliability eoefficients given in the ensuing tables are too low, if anything, rather than too high.

Results

The results of the *individual* experiments are summarized in Table 1. With both 2 mm and 3 mm stimulators, the correlations between successive periods are .85 or better in two-thirds of the cases and above .75 in the remainder. These represent a combined measurement of the stability of sensitivity and the consistency of the subjects in reporting, somewhat reduced by the artifacts previously mentioned. On the whole, then, the performance of these totally untrained subjects seems highly satisfactory. Indeed, reliability coefficients of .93 and .96 could scarcely be exceeded by the most highly-trained observer.

The coefficients for the *group* experiments are shown in Table 2, where all possible intercorrelations are given. Since the experimenters in this work were as untrained as the subjects, a somewhat lower general level of performance is to be expected. But the peak performances compare quite favorably with those obtained on the individual basis. Six subjects reached .85 or better and six more .75 or better. The remaining three were decidedly inferior, two of them barely getting above .65 and the other giving little better than a random performance. Naturally, the correlations

TABLE 1

INDIVIDUAL EXPERIMENTS

3 mm Diam. Stimulator—17° C.

Subj. Init.	Correl. × 100	Distribution of Scores				
		0–1	2–3–4	5–6–7	8–9–10	11–12
FLB	93	31 29	7 7	4 13	8 1	0 0
SBP	91	14 16	12 13	18 18	5 3	1 0
JLB	87	3 1	11 7	9 14	11 14	16 14
HFJ	85	2 2	9 5	9 17	9 5	21 21
MBR	81	0 0	7 2	16 6	10 19	17 23
FBS	77	5 1	6 2	9 3	8 6	22 38
2 mm Diam. Stimulator—17° C.						
HLR	96	24 27	8 7	12 12	5 4	1 0
HWT (1)	86	5 1	10 8	9 20	11 10	15 11
HWT (2)	89	3 2	6 2	13 19	14 10	14 17
RMS	85	6 5	12 8	24 16	7 13	1 8
HTSH	80	21 25	12 13	10 9	7 3	0 0
JGH	79	11 14	18 14	15 21	6 1	0 0

fall off somewhat with two and three days intervals, but are still at a respectable level.

The distribution of scores for the group experiments (Table 3) indicates considerable shifting from period to period. As previously noted, it is difficult to determine the cause in most instances. With subject TR, however, there is a continual downward trend, which might be taken as indication of a

change in sensitivity of the area as a whole. It should be noted that this subject showed fairly high correlations when only successive periods were compared, in spite of marked negative adaptation effects during the course of every period. Improvement of discrimination with practice is clearly shown by subjects KK, AED and JH, as evidenced by the increasing grouping around the category levels. Note that in KK's final period the intermediate scores 2–3–4 and 8–9–10 disappeared entirely.

TABLE 2

CORRELATIONS × 100—GROUP EXPERIMENTS

3 mm Diam. Stimulator—17° C.

Interval	One Hour		One Day					Two Days				Three Days			
Periods / Subj. Init.	1A 1B	4A 4B	1A 2	1B 2	2 3	3 4A	3 4B	1A 3	1B 3	2 4A	2 4B	1A 4A	1A 4B	1B 4A	1B 4B
KK....	87	84	88	92	89	89	89	83	89	77	81	76	75	77	73
AED...	74	94	70	84	85	90	87	70	78	86	90	63	67	80	82
WB....	79		87	81	89	80		80	85	85		85		79	
AMC...	96		91	92	91			82	92						
DE....	77		78	74	80			73	66						
MLF...	78		77	70	77			78	72						
JH.....	76		72	89	85			77	84						
ADP...	75	82	64	72						91	62				
TR.....	78		82	81	63			62	52						
RBS....	48		19	00	39			51	20						
CD....					54	66				57					
FSF....					75	72				57					
WK....					70	62				66					
REB...					84										
JED....		93													

Possibly more important than the individual correlations is the general consistency over several days' time as shown in Fig. 2, where average scores (rows) are plotted against individual scores (columns). An example may make this somewhat unorthodox graph clearer. With subject KK, there were 8 squares with an average score of '6' for the entire session. For each of these squares, we have available 6 individual scores, or 48 in all. Examination of the chart shows that 42 of these 48 scores were identical with the average score or within one point of it. Among all of the

better subjects, the clustering of values along the diagonal is notable. The mean deviation for any average score is typically less than one step interval and 'wild shots' are comparatively

TABLE 3

Distribution of Scores—Group Experiments

3 mm Diam. Stimulator—17° C.

Subj. Init.	Period	0–1	2–3–4	5–6–7	8–9–10	11–12
KK	1A	2	8	13	14	13
	1B	3	10	15	11	11
	2	0	9	15	12	14
	3	1	4	24	5	16
	4A	1	3	26	9	11
	4B	2	0	28	0	20
AED	1A	0	2	14	20	14
	1B	0	2	15	17	16
	2	0	2	20	8	20
	3	0	0	16	11	23
	4A	0	0	17	9	24
	4B	0	0	19	3	28
WB	1A	3	12	12	12	11
	1B	2	13	7	16	12
	2	4	11	10	11	14
	3	5	7	11	15	12
	4A	2	7	12	17	12
AMC	1A	9	4	9	12	16
	1B	9	4	9	10	18
	2	8	5	9	7	21
	3	10	2	4	12	22
DE	1A	3	5	13	12	17
	1B	0	7	9	14	20
	2	5	6	4	11	24
	3	1	5	8	10	26
MLF	1A	6	9	10	13	12
	1B	2	6	12	18	12
	2	3	10	4	18	15
	3	3	13	10	15	9
JH	1A	3	8	11	18	10
	1B	2	10	12	14	12
	2	5	6	19	6	14
	3	4	6	21	4	15
ADP (1)	1A	0	6	8	13	23
	1B	0	3	6	16	25
	2	0	0	5	14	31
ADP (2)	3	3	4	10	17	16
	5A	1	7	11	17	14
	5B	2	6	7	17	18

TABLE 3 (*continued*)

Subj. Init.	Period	0–1	2–3–4	5–6–7	8–9–10	11–12
TR	1A	3	9	24	10	4
	1B	7	11	23	7	2
	2	12	17	17	4	0
	3	16	17	15	1	1
RBS	1A	0	1	4	27	18
	1B	0	0	2	16	32
	2	0	0	6	18	26
	3	0	1	7	13	29
CD	2	0	17	30	2	1
	3	0	9	36	5	0
	4A	0	1	29	18	2
FSF	2	0	6	19	18	7
	3	0	8	23	13	6
	4A	0	2	27	18	3
WK	2	1	11	14	16	8
	3	3	8	23	14	2
	4A	0	6	23	16	5
REB	2	6	19	12	11	2
	3	9	11	21	5	4
JED	5A	4	4	11	13	18
	5B	3	6	18	9	14

rare. This makes the average score for a square a highly meaningful value.

The significance of these average scores is likewise indicated by the maps in Fig. 3. The number in each square of the map is the average score from all available data. It must be remembered that these squares were not stimulated in a fixed order. All of the black squares were stimulated first; then all of the white squares, each group in a random fashion. Yet, when the two sets of data are put together, a striking *orderliness* is revealed. Squares of high sensitivity are not scattered irregularly through the map, but typically occur in fairly large clusters. Frequently, the tapering-off to lower and lower levels can be traced. Even those numbers which occur very infrequently tend to be grouped. Notice, for instance, the vertical string of 6's in the map for AED or the block of 3's, 4's and 5's in that of KK. Possibly the most significant feature of these maps is the indication of lawfulness

even when the reliability coefficients are fairly low. The average scores from a sufficient number of trials may be meaningful in such cases, although it is impossible to establish high consistency when period is compared with period.

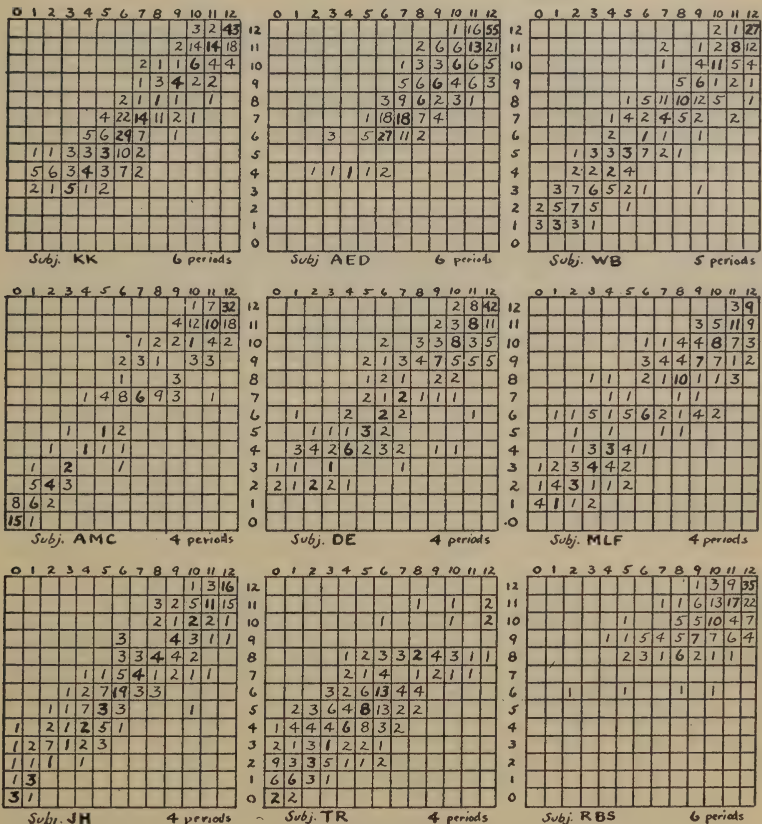


FIG. 2. Average against individual scores.

As a matter of interest, maps derived from the individual experiments with 2 mm diam. stimulator (average of only two periods) are shown in Fig. 4. While the orderliness of these maps is much less striking than in Fig. 3, this may in part be an artifact of the smaller number of cases used in computing the average scores.

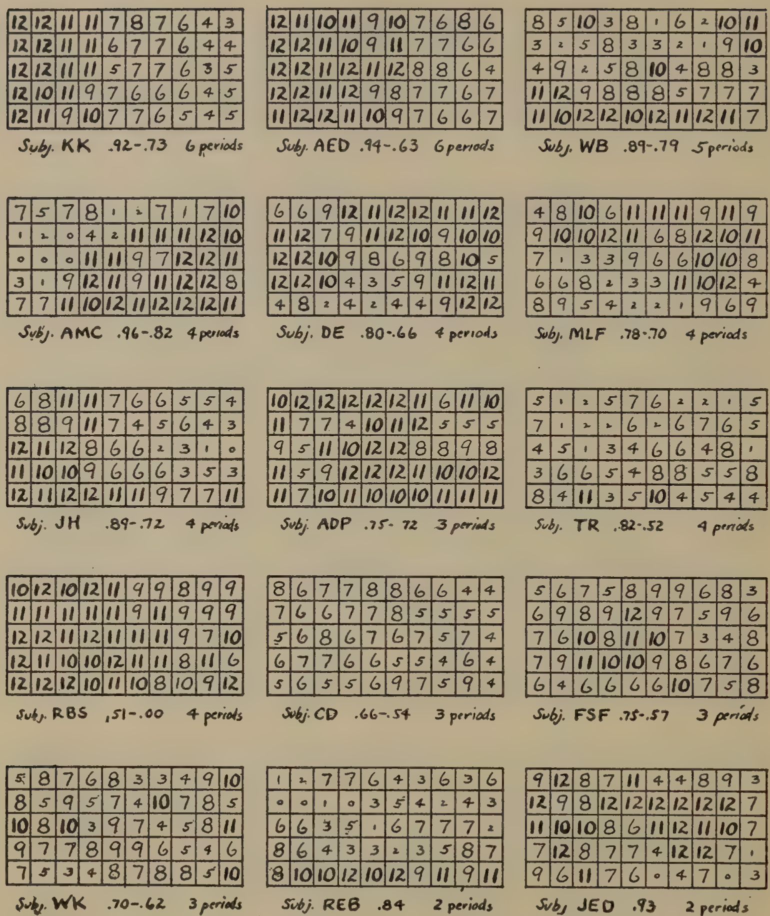


FIG. 3. Maps of average scores—3 mm stimulator.

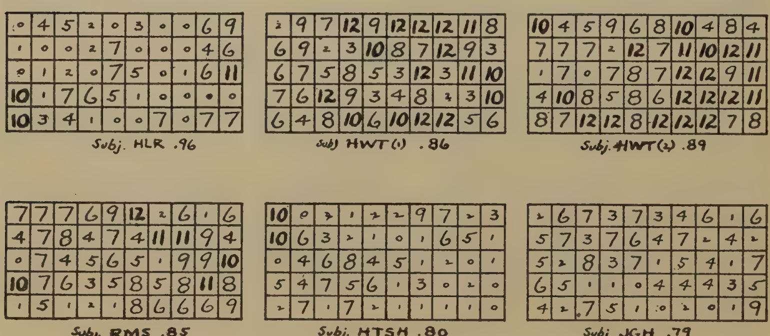


FIG. 4. Maps of average scores—2 mm stimulator.

Significance of Results

From the foregoing data, it becomes evident that seriatim mapping provides a promising research technique for the *quantitative* study of the effects of low temperature stimulation. Four essential steps suggest themselves:

1. Preliminary tests to determine reliability coefficients. There is some indication that good subjects improve rather rapidly; whereas initially-poor subjects show little benefit from training.

2. Sufficient training to establish reasonably consistent standards of judgment.

3. Establishment of a proper scheme of categories. Too many numbers are confusing to the untrained subject. Too few of them result in an artificial grouping of scores around the category levels as discrimination improves.

4. Sufficient number of tests for the reliable determination of average scores. Statistically, the reliability increases with the square root of the number of cases. Practically, a number of other factors must be taken into consideration.

Some of the *problems* which might profitably be investigated with such a technique include:

1. The effect of stimulator size. A rough comparison of the 1 mm maps from the Dallenbach study with the 2 mm and 3 mm maps from the present investigation shows interesting possibilities.

2. The effect of stimulator form. It has already been shown that the shape of stimulator, as such, markedly affects adaptation time [4] and may also alter the original intensity of experience.

3. The effect of stimulating temperature, especially at levels close to the limen.

4. Shifts in sensitivity over an extended period of time. There is some indication that extremely strong and extremely weak squares tend to remain stable, while those in the intermediate levels do most of the shifting.

[4] Jenkins, W. L., Studies in thermal sensitivity: 2. Adaptation with a series of small rectangular stimulators, *J. Exper. Psychol.*, 1938, 22, 84–89; 3. Adaptation with a series of small annular stimulators, *J. Exper. Psychol.*, 1938, 22, 164–177.

The outcome of a systematic program of this kind should be a clearer understanding of the *necessary* scheme of cold-receptor distribution. The usefulness of the concept of discrete 'cold spots' (as marked by a very small stimulator) is already under suspicion. It should be possible to demonstrate conclusively whether the actual distribution of cold sensitivity can be explained in terms of such 'cold spots' at all. The present results suggest as a more fruitful hypothesis the assumption of a large number of minute receptors, with great differences in the density of their distribution. Thus, peaks of sensitivity would appear where the density is the greatest and valleys where it is least . . . the so-called 'cold spots' being merely artifacts of the method of stimulation employed, particularly stimulator size and temperature. Such an assumption would also change materially the course of our search for the receptors involved.

Any such speculation, however, should wait upon the results of much further research—for which the seriatim technique seems admirably adapted.

Summary

When the results of Dallenbach's study with trained observers (four single cold-mappings on different days with a 1 mm diam. stimulator at 8–9° C.) are replotted in terms of composite maps showing the number of times each mm square was positively reported, no high consistency is established. Either the mapping must be faulty, or the 'cold spots' are inconstant. If the concept of discrete 'cold spots' is discarded, the composite maps may be viewed as an orderly scheme of peaks and valleys of sensitivity, from which consistency in mapping may be inferred, but is not proved. This ambiguous outcome points to the need for a seriatim technique for the independent determination of reliability.

In the present study, seriatim cold-mapping was carried out upon totally untrained subjects with 2 mm and 3 mm stimulators at 17° C. Each square of a checkerboard pattern was stimulated 6 times during a period; repeated mappings

being made the same day and also for several days in succession. Three indications of consistency are found: (1) reliability coefficients of .85 or better are obtained in the majority of cases when successive periods are compared; (2) charts of average against individual scores show low mean deviations and few 'wild shots'; (3) maps derived from average scores indicate a high degree of orderliness, even when the reliability coefficients are low. It is concluded that seriatim cold-mapping is a promising research technique. Certain avenues for further investigation are suggested.

(Manuscript received August 30, 1938)

Reprinted from the Journal of Experimental Psychology,
Vol. 25, No. 3, September, 1939
Printed in U. S. A.

STUDIES IN THERMAL SENSITIVITY: 11. EFFECTS OF STIMULATOR SIZE IN SERIATIM COLD-MAPPING

BY WILLIAM LEROY JENKINS
Lehigh University

In a previous article,[1] it was shown that highly reliable results can be obtained from untrained subjects in seriatim cold-mapping with 2 mm and 3 mm diam. stimulators. In the present study, a wider range of stimulator sizes has been used upon a larger number of untrained subjects, to secure some preliminary data upon the influence of stimulus area.

Procedure.—Briefly, the seriatim method as here carried out involves repeated stimulation of a suitable checkerboard pattern so that a 'score' is obtained by summing the reports. Four categories are employed: 0 for neutral, 1 for weak, 2 for medium and 3 for strong cold. (Warm, if experienced, is simply reported as 'W'.) Six rounds of stimulation constitute a series. Thus the maximum score is 18 and the minimum 0. The scores are merely pseudoquantitative, since the categories 1, 2 and 3 are ordinal and not cardinal numbers. Defense of the statistical procedure and an analysis of the probable artifacts in mapping have been presented in the previous paper.[2]

All of the experiments here reported were performed by the group method, three students to a group, one acting as experimenter and the other two alternating as subject and recorder. To equalize the effects of adaptation as far as possible, a uniform time of 3½ min. per round was maintained; so that approximately 7 min. elapsed between successive stimulations of the same squares. A practically random order of stimulation was employed, and the general sequence continually shifted so that the subject (who worked with his eyes closed) could not memorize a pattern of reports.

Five stimulator sizes were used: 3 × 3 mm, 3 × 1½ mm, 2 × 2 mm, 1½ × 1½, and a circular stimulator 1 mm in diam. The last-named approaches the limit for accurate manual placement.

Results

Because the same students were not available for all of the experiments, the following results are not strictly comparable on a quantitative basis. However, the major implications of the data are sufficiently clear.

[1] Jenkins, W. L., Studies in thermal sensitivity: 9. The reliability of seriatim cold-mapping with untrained subjects. *J. Exper. Psychol.*, 1939, *24*, 278–293.

[2] *Op. cit.*, 282–284.

1. *Reliability Coefficients.*—Table 1 summarizes the reliability coefficients obtained when the scores from two successive series of 6 rounds each are compared. With all stimulator sizes, at least 75 percent of the subjects reached .70 and approximately half of them .80—with a few exceptional individuals making .90 or higher. Where low reliability was obtained, the chief cause seemed to be the failure to discriminate a sufficient range of experiences. The few subjects below .60 typically employed only two categories of report. Occasionally, a subject who could do well with the largest size (3 × 3 mm) would fail miserably with the smaller ones.

TABLE 1

RELIABILITY COEFFICIENTS *

Stim. Size	Below .60	.60 .64	.65 .69	.70 .74	.75 .79	.80 .84	.85 .89	.90 .94	.95 1.00	Total Cases
3×3......	0	2	3	5	2	7	8	3	1	*31*
3×1½.....	1	0	1	3	5	3	6	1	0	*20*
2×2......	11	4	6	8	14	13	20	6	0	*82*
1½×1½....	2	0	3	2	2	5	6	1	0	*21*
1 mm.....	0	2	3	2	6	2	4	5	4	*28*

* Figures in the table indicate the number of cases within the given step interval.

Virtually all subjects attained somewhat better results with the 3 × 3 mm than with the 2 × 2 mm stimulator.

Possibly the most surprising result is the high percentage of coefficients over .90 with the 1 mm diam. It might be thought that this was due to reporting on an all-or-none basis. Such was not the case. Although the '3' category virtually disappears at this level and the number of consistent zeros is large, 9 of the 11 subjects above .90 used both '1' and '2' consistently. It will be recalled that the 1 mm diam. is the size of stimulator used by Dallenbach in his study of maps repeated on four different days.[3] Since some of our *untrained* subjects were able to reach .90 or better, it is reasonable to assume that Dallenbach's *trained* observers would do equally well. This supports our previous contention [4] that the discrepancies in the four mappings were indications of actual gradations of

[3] Dallenbach, K. M., The temperature spots and end-organs. *Amer. J. Psychol.*, 1927, 39, 402–427.

[4] *Op. cit.*, 280.

sensitivity within the area (hills and valleys) and not of poor work on the part of experimenters and observers.

In the balance of this article, only the data from subjects with reliability coefficients of .80 and over will be cited. This criterion is probably unnecessarily stringent, but will avoid any possibility that the conclusions are distorted by purely random reporting.

2. *Distributions of Scores.*—Table 2 shows the *average* distribution of scores for each of the five stimulator sizes, together with the individual distributions for the subjects having the maximum and minimum gross scores in each case. Only subjects with reliability coefficients of .80 and over are included. The grouping: 0–1, 2–3–4, 5–6–7, etc. represents roughly the four category levels and the three intermediate positions.

TABLE 2

Distributions of Scores

Stim. Size		0–1	2–3–4	5–6–7	8–9–10	11–12–13	14–15–16	17–18
3×3	(Av.)....	*7%*	*13%*	*21%*	*20%*	*18%*	*15%*	*6%*
	(Min.)...	24	22	20	14	10	6	4
	(Max.) ..	0	2	10	18	28	30	12
3×1½	(Av.)....	*12%*	*18%*	*18%*	*16%*	*16%*	*11%*	*9%*
	(Min.)...	48	26	26	0	0	0	0
	(Max.) ..	0	0	4	6	16	32	42
2×2	(Av.)....	*13%*	*16%*	*23%*	*19%*	*15%*	*9%*	*5%*
	(Min.)...	66	22	8	0	4	0	0
	(Max.) ..	0	0	4	20	22	26	28
1½×1½	(Av.)....	*15%*	*25%*	*25%*	*15%*	*11%*	*7%*	*2%*
	(Min.)...	47	25	28	0	0	0	0
	(Max.) ..	2	3	12	20	26	27	10
1 mm	(Av.)....	*59%*	*16%*	*12%*	*5%*	*5%*	*2%*	*1%*
	(Min.)...	80	5	15	0	0	0	0
	(Max.) ..	22	20	10	15	20	8	5

This *average* distribution of scores is a purely mathematical construct. The actual distributions varied between the maximum and minimum levels shown, so that the average can scarcely be considered even as typical. However, two chief points do stand out: (1) With stimulators 1½ × 1½ and larger, the distributions are all very similar and there is con-

siderable overlapping. (2) Between the 1½ × 1½ and 1 mm diam. sizes, there appears to be a sharp break which is likewise reflected in the maximum and minimum scores.

3. *Average Scores.*—Table 3 shows the average scores for subjects with reliability coefficients of .80 and over, and the mean score for each stimulator size. The median score for each size agrees closely with the mean, except in the case of the 1 mm diam. stimulator where a few high values cause an

TABLE 3

Average Scores *

Average Scores	Stimulator Sizes and Areas				
	3 × 3 9.0 mm.^{2}	3 × 1½ 4.5 mm.^{2}	2 × 2 4.0 mm.^{2}	1½ × 1½ 2.25 mm.^{2}	1 mm. 0.8 mm.^{2}
0–1	—	—	—	—	—
1–2	—	—	1	—	8
2–3	—	1	1	1	1
3–4	—	—	2	—	2
4–5	—	1	3	3	2
5–6	1	—	6	1	1
6–7	4	1	3	2	1
7–8	1	2	5	3	—
8–9	5	1	4	1	—
9–10	2	2	4	—	—
10–11	2	1	7	—	—
11–12	2	—	1	1	—
12–13	2	—	—	—	—
13–14	—	—	2	—	—
14–15	—	1	—	—	—
Total	19	10	39	12	15
Mean score for group	*9.0*	*8.1*	*7.7*	*6.4*	*2.8*

* Figures in the table indicate number of cases within the given step intervals.

upward distortion of the mean. Taking this into account, it appears that the averages for the 1 mm diam. and 1½ × 1½ mm sizes are closely proportional to the relative stimulus areas. Taken at its face value, this would indicate a roughly linear summation effect within this range. Above this level, however, a plotted curve would flatten out rapidly. A change from 4.5 mm^2 to 9.0 mm^2, for example, brings about an increase of less than one unit in the average score. Again, it must be emphasized that these 'average' values are not quantitatively significant but may be taken only as general indicators of a trend.

4. *Maps*.—The general appearance of the maps derived from subjects showing reliability coefficients of .80 and over does not differ significantly from those shown in the earlier study.[5] With all stimulator sizes, there is a tendency for similar levels of sensitivity to be grouped. Even with the 1 mm diam., the picture is typically one of hills and valleys, like a topographic map, and not of a random arrangement of scattered 'cold spots.' As indicated by the wide range of average scores, the general level of reports varies markedly from subject to subject, and the appearance of the maps changes correspondingly.

Discussion

The foregoing should be considered as no more than a preliminary attack on the problem. It is evident that simple linear summation does not occur with all stimulator sizes. It may be an oversimplification to assume that it holds even in the lowest range. The chief value of the study lies in showing the feasibility of the seriatim technique for this purpose. Since good subjects can report consistently over a range of stimulus sizes, as shown by the high reliability coefficients, it should now be possible to map the same checkerboard area with two different sizes. The comparative scores will then give a basis for evaluating summation effects with supra-liminal stimulation.

Summary

Seriatim cold-mapping with a range of five stimulator sizes, the smallest 1 mm diam., gives satisfactory reliability coefficients (.80 or over) with all sizes in the majority of untrained subjects. Average scores indicate roughly a linear summation effect at the lowest levels, but the curve flattens out rapidly with the larger sizes. Even with the 1 mm diam. stimulator, the maps tend to show grouping of similar scores into hills and valleys, and not a random arrangement of separate cold spots. By mapping the same area with two different sizes of stimulator, it should be possible to attack the problem of summation with supra-liminal stimuli.

(Manuscript received February 17, 1939)

[5] *Op. cit.*, 290.

Reprinted from the JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
Vol. 25, No. 4, October, 1939
Printed in U. S. A.

STUDIES IN THERMAL SENSITIVITY: 12 PART-WHOLE RELATIONS IN SERIATIM COLD-MAPPING

BY WILLIAM LEROY JENKINS

Lehigh University

In a preliminary study of seriatim cold-mapping with stimulators of different sizes,[1] it has been shown that simple summation is not the rule; *i.e.*, that doubling the stimulus area does not yield twice as large a seriatim score. In the present investigation, to determine more accurately the relation between stimulus area and seriatim scores, the same cutaneous regions have been mapped alternately with two stimulators, one being one-half or one-third the area of the other.

PROCEDURE

The method of seriatim mapping was similar to that employed in the preceding study.[2] Untrained student subjects worked in groups of three, one acting as experimenter, the other two alternating as subject and recorder. A checkerboard pattern of suitable squares or rectangles, stamped on the volar surface of the left forearm, was mapped 12 times—6 rounds with each size.[3] The report categories were: 0 for neutral, 1 for weak, 2 for medium and 3 for strong cold. Inlet temperature was maintained at 17° C. in all experiments.

Figure 1 shows some sample data from one row of squares (10 out of 50) in the series employing 3 × 3 and 3 × 6 stimulators alternately. The upper blocks give the successive reports from 3 × 6 stimulation; the lower those from 3 × 3 stimulation. Subjects do not always respond with exactly the same number to repeated stimulation of the same square. Occasionally, a completely erratic report will appear—possibly the result of misplacement of the stimulator by the experimenter or of inattention on the part of the subject. The simple sum of the numbers is always taken as the *seriatim*

[1] Jenkins, W. L., Studies in thermal sensitivity. 11. Effects of stimulator size in seriatim cold-mapping, *J. exper. Psychol.*, 1939, **25**, 302–306.

[2] *Op. cit.*

[3] Consider the 3 × 3—3 × 6 series as an example: The sequence was: 3 × 3 on first subject; 3 × 3 on second subject; 3 × 6 on first; 3 × 6 on second, etc. until 6 rounds with each size had been mapped on each subject. A minimum time of 3½ min. per round was rigidly maintained, to equalize adaptation effects. The stimulation followed approximately a random order to avoid any possibility that the subject could memorize a sequence of reports.

3 × 6	2	0	3	3	1
	3	0	3	3	1
	2	0	3	3	1
	3	0	3	3	1
	3	1	3	2	1
	2	0	3	2	1

3 × 3	2	2	1	0	2	3	2	3	1	2
	1	1	0	1	3	3	3	3	2	1
	1	2	1	1	2	2	3	2	2	0
	2	3	2	1	2	3	3	2	1	0
	2	2	0	0	2	2	2	3	1	0
	3	3	2	1	3	3	3	2	1	1

FIG. 1. Excerpt from typical data sheets.

score, however, without any attempt at arbitrary correction. This summing of ordinal numbers as if they were cardinal is theoretically reprehensible, but appears to be empirically justified in this case by the consistent relations among the scores thus computed. In practice, 0, 1, 2 and 3 must approach a linear series for such relations to hold.

Results

Table 1 gives both the reliability coefficients from repetition with the same size and the coefficients of correlation between whole scores and mean part scores when different sizes

TABLE 1

From Repetition with the Same Size

Stim. Size	.00 .59	.60 .64	.65 .69	.70 .74	.75 .79	.80 .84	.85 .89	.90 .94	.95 1.00	Total Cases	
3 × 3	0	2	3	5	2	7	8	3	1	31	
3 × 1½	1	0	1	3	5	3	6	1	0	20	52%
2 × 2	11	4	6	8	14	13	20	6	0	82	over
1½ × 1½	2	0	3	2	2	5	6	1	0	21	.80
1 mm.	0	2	3	2	6	2	4	5	4	28	
Between Total Scores and Average Part Scores											
3 × 6 3 × 3	1	2	2	6	5	8	1	7	0	32	
3 × 3 3 × 1½	2	2	1	5	5	6	4	2	1	28	50%
2 × 4 2 × 2	3	1	2	2	3	5	6	4	0	26	over
3 × 1½ 1½ × 1½	1	0	3	3	2	1	6	2	0	18	.80
2 × 2 2 × 1	4	1	0	1	4	3	4	3	0	20	
2 × 1 1 mm.	7	1	4	2	5	6	1	2	0	28	
3 × 4½ 3 × 1½	4	0	0	1	1	1	4	1	0	12	
2 × 6 2 × 2	2	1	4	3	3	8	8	3	0	31	

Figures in the table indicate the number of cases within the given step interval.

are used alternately. In both instances, these are .80 or higher in approximately 50 percent of the cases, indicating some extremely orderly relation between whole and part scores. To avoid any possibility that the results might be distorted by random reporting, only the data from those subjects with correlations of .80 and over have been used in the ensuing analyses. (As a matter of fact, almost identical curves are obtained by using all the data indiscriminately.)

The Gross Relation to Stimulator Size.—By combining the data for all reliable subjects, we obtain Fig. 2, in which average

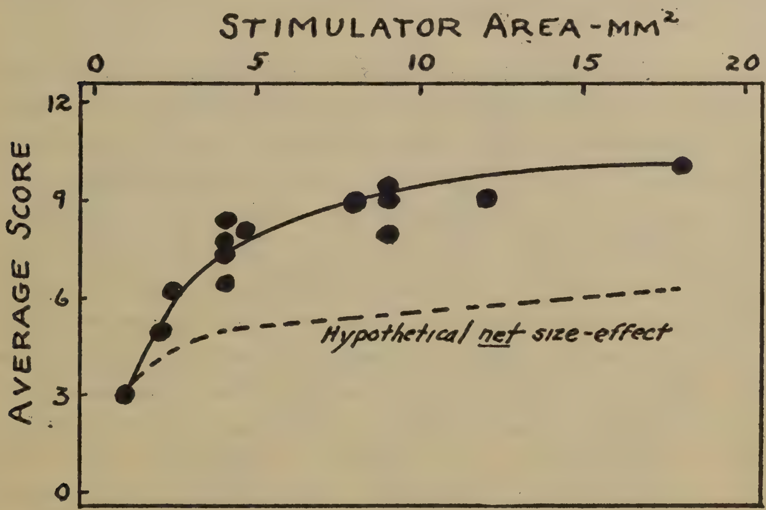


FIG. 2

scores are plotted against stimulator area, the circles indicating average scores for different groups of 10 or more subjects, and the solid line approximately a best-fitting curve. This is not intended as a quantitative representation, but sets the stage for the next question:

To what extent is this curve a function of stimulator area as such? The answer must be approached indirectly.

The Hypothetical Net Size-Effect.—Table 2 shows whole and part average scores from alternate mapping—selecting only those cases where a whole score can be matched with two *identical* part scores. The percentage increase of whole over

TABLE 2

Part Size	Whole Size	Mean Part Score	Mean Whole Score	Increase: Whole over Part
Square-to-Rectangle				
3 × 3	3 × 6	9.1	9.8	8%
2 × 2	2 × 4	8.0	8.2	1%
1½ × 1½	3 × 1½	3.7	4.1	11%
Rectangle-to-Square				
3 × 1½	3 × 3	5.0	6.0	20%
2 × 1	2 × 2	2.8	3.4	21%

part score in such instances represents the effect of stimulator area as such, divorced from any possible influence of part score differences. The increases range from 1 percent to 21 percent.[4] By applying these percentages successively, we can construct a hypothetical curve which should represent the *net* influence of stimulator area. This is shown as the dotted line in Fig. 2. A liberal allowance has been made for the increase from 1 to 2 mm^2, because the average score figures at the smallest size are so overweighted with double-zeros as to make direct comparison impossible. The hypothetical net size-effect curve [5] clearly falls far below the solid line in Fig. 2 which represents the gross effect of stimulator size. The implication is plain: Some factor other than stimulator size as such must operate. This factor must be related to part score *differences*.

The Basic Curve.—As a preliminary step toward identifying this function, we must abandon averages and investigate the relation between individual whole and part scores. The upper left-hand graph of Fig. 3 results from plotting individual

[4] The square-to-rectangle increase is consistently less than that for rectangle-to-square, which implies that rectangular stimulation is 'weaker.' The effect of stimulator form and stimulator orientation are topics for further investigation.

[5] Theoretically, by mapping an area successively with larger and larger stimulators and selecting the ever-diminishing number of cases showing identical part scores at all size levels (indicating uniform sensitivity), it would be possible to develop the net size-effect curve empirically. Since we are interested in the curve only provisionally, this time-consuming task has not been undertaken.

whole scores against individual part scores, selecting those cases where *identical* parts can be matched with whole scores. For this purpose, all such cases from the six whole-half comparisons have been combined, neglecting the variations in percentage of increase shown in Table 2. The circles indicate the medians of whole scores, with open circles where these are

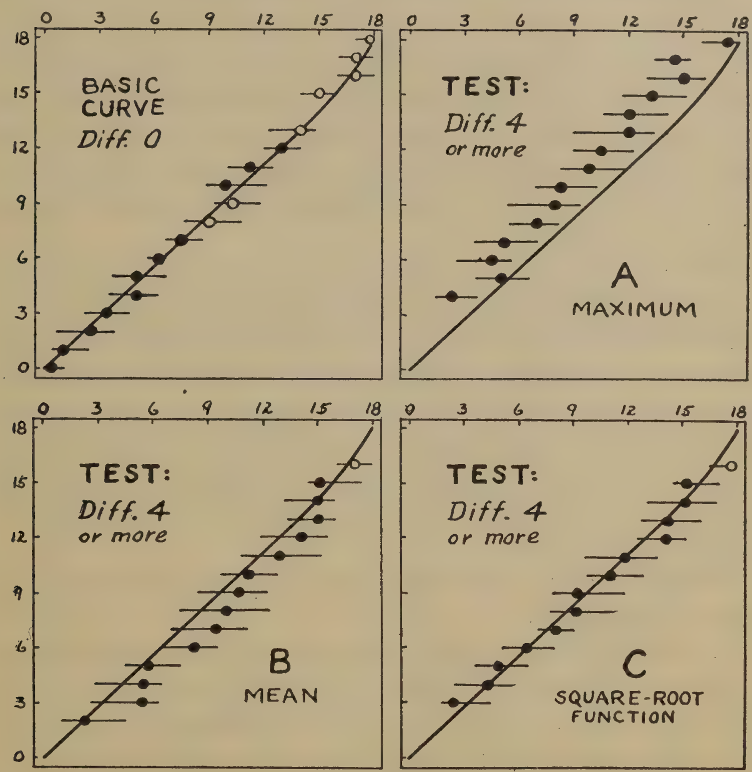


FIG. 3. Preliminary tests. Ordinates: functions of part scores. Abscissæ: whole scores. Circles: medians of whole scores. Lines: interquartile ranges.

based on less than 10 cases. Horizontal lines indicate the interquartile ranges.[6]

[6] The small size of the quartile deviation—less than $1\frac{1}{2}$ units generally—provides empirical justification for the method of determining scores by adding the category numbers reported. Although there are wide individual differences in score levels—although each subject sets his own standards for what he will call 0, 1, 2 and 3—there is evidently a common basis in the *relationship* between 0 and 1, 1 and 2, and 2 and 3 as they are employed by most subjects.

The heavy line in this graph represents approximately a best-fitting curve. The slope of the linear portion indicates an average 8 percent increase in whole over part scores, which agrees substantially with the smaller average score increases in Table 2 and provides further justification for our hypothetical net size-effect curve of Fig. 2.

Preliminary Tests.—The basic curve has been plotted only from cases where the part scores are identical. The next step is to discover a function of the part scores which provides a good fit when the components are *non-identical.* Since the true relation will be revealed more sharply by cases with large differences, only those where part scores differ by 4 or more units are used in the preliminary tests shown in the remaining graphs of Fig. 3.

The *maximum* (*A*) of part scores is clearly not the correct function. Whole scores are lower than maximum part scores. The *mean* (*B*) is somewhat better, but whole scores are higher than mean part scores.[7] The *square-root function* (*C*) provides a fit to the basic curve which is approximately as good as that of the identical score medians from which it was derived. This function is the square root of the mean of the squares of the part scores (mathematically similar to the formula for standard deviation).

Critical Checks.—Two checks for this function immediately suggest themselves:

1. Is the fit to the basic curve good with all magnitudes of part score differences? Figure 4 provides the answer. With differences of 1–2–3 and 4–5–6, the coincidence is excellent. With differences of 7 or more, six of the medians touch the curve, while four are about one score point higher. It is dangerous to explain away such discrepancies as 'experimental errors,' because further investigation may show that some modification of the function is necessary. Nevertheless, the following opportunity for artifacts must be noted: Where large differences in part scores exist, it means that the level of sensitivity is changing sharply. Consequently, even slight

[7] In any event, the mean would be a poor choice since it gives no chance for an influence of part score differences.

misplacements of the stimulator will lead to erratic reports. In a region where there is a *general* predominance of high sensitivity, it is conceivable that misplacements into regions of higher sensitivity might occur more frequently than into regions of lower sensitivity. This would automatically make whole scores a trifle higher than the basic curve calls for.

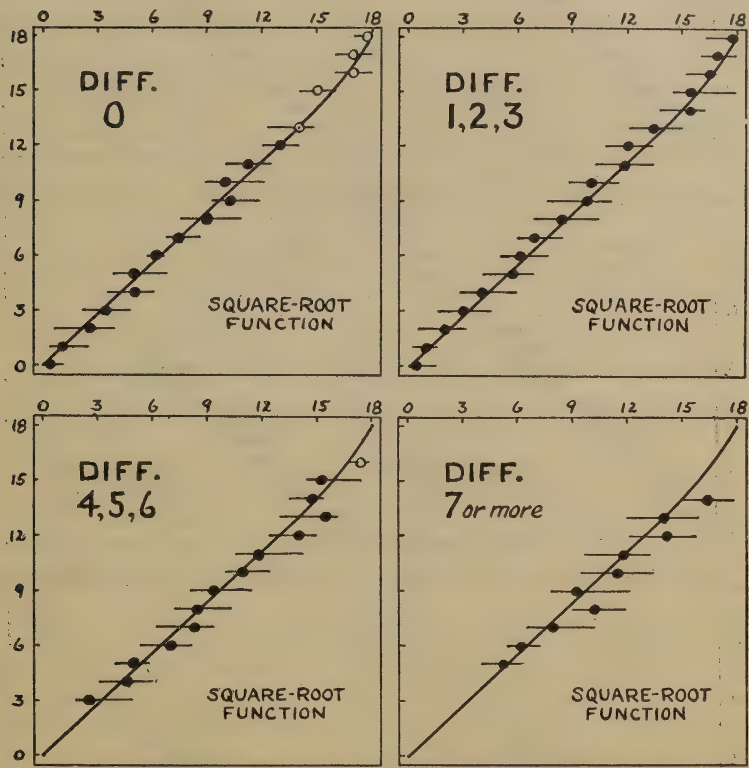


FIG. 4. Square-root function applied to various part score differences. Ordinates: square-root function of part scores. Abscissæ: whole scores. Circles: medians of whole scores. Lines: interquartile ranges.

2. Is the fit to the basic curve good with all size comparisons? Figure 5 shows that it is. Even the slight departures to the right or left of the basic curve agree with the varying percentage increases shown in Table 2, which was based on identical part scores. Thus, the medians for $3 \times 4\frac{1}{2}$

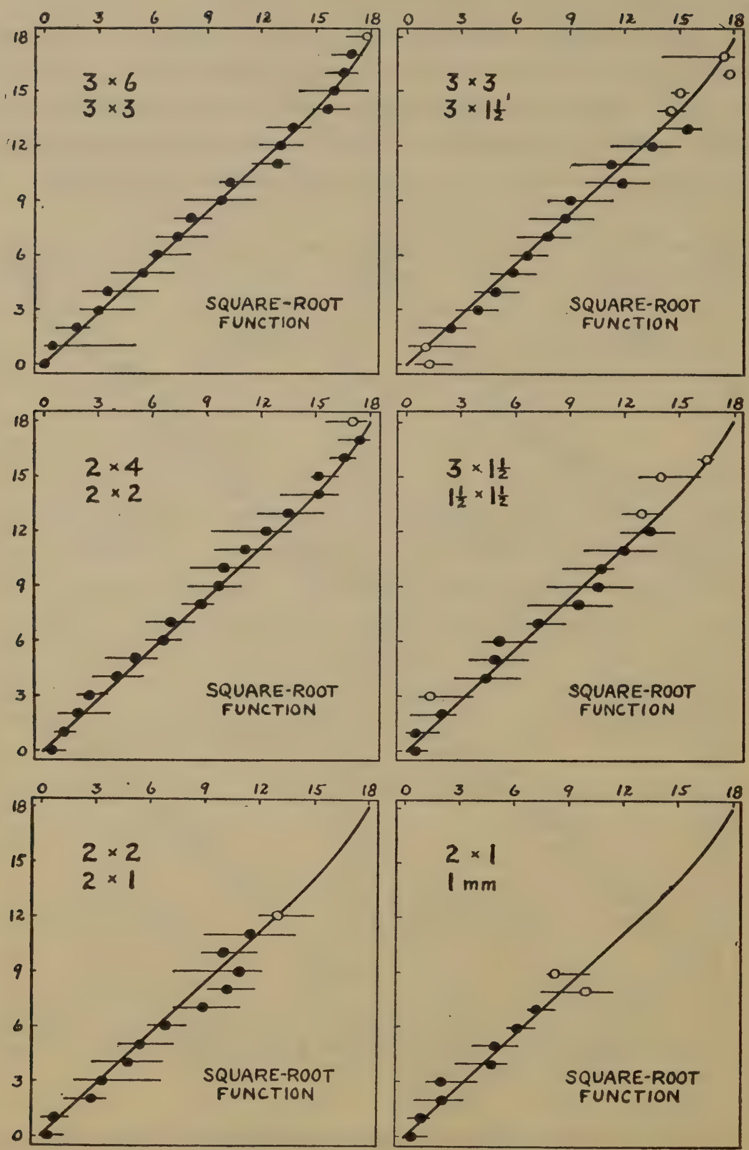


FIG. 5. Square-root function applied to whole-half comparisons. Ordinates: square-root function of part scores. Abscissæ: whole scores. Circles: medians of whole scores. Lines: interquartile ranges.

$- 3 \times 1\frac{1}{2}$ and $2 \times 2 - 2 \times 1$ (20% and 21% in Table 2) veer to the right, while the medians for $2 \times 4 - 2 \times 2$ (1%) are slightly to the left, and those of $3 \times 6 - 3 \times 3$ (11%) are almost perfectly centered. Confirmation is furnished by Fig. 6 where two whole-third comparisons are presented. We have no basic curve for whole-third relations, because cases with three identical components are rare. However, the fit to the whole-half basic curve is very good, especially for the $2 \times 6 - 2 \times 2$ series.

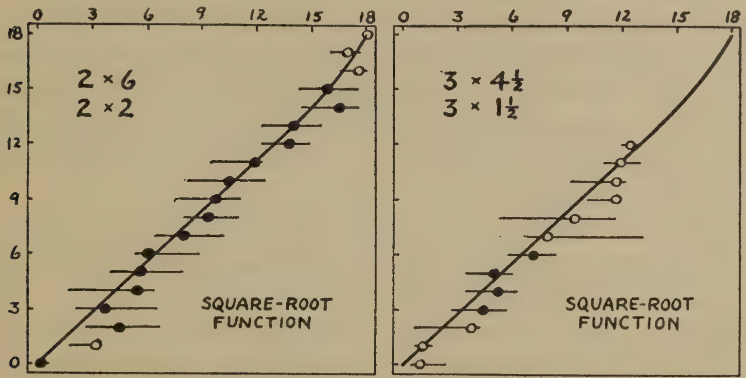


FIG. 6. Square-root function applied to whole-third comparisons. Ordinates: square-root function of part scores. Abscissæ: whole scores. Circles: medians of whole scores. Lines: interquartile ranges.

From the results of these critical checks, it appears that the square-root function holds within the range of stimulator sizes investigated. If not an exact mathematical statement of the relation between whole and part scores (when size-effect as such is excluded), it is certainly a satisfactory first approximation.

Discussion

'*Cold Spots' Inadequate.*—Since the pioneer studies of Blix, Goldscheider and Donaldson, it has been generally held that cutaneous sensitivity is punctiform. In a strictly descriptive sense, this merely means that sensory 'spots' can be mapped on the skin with small stimulators. However, if such 'spots' are to have any intrinsic significance—if they are

to be more than incidental artifacts of a particular method of mapping—they must represent ultimate units corresponding to receptors. It has been generally assumed that the 'spots' *do* correspond to receptors. Thus, we have the von Frey hypothesis that 'cold spots' are Krause end-bulbs—a view which Strughold, Bazett and others have attempted to support with histological evidence. Even those who dispute the von Frey theory have generally sought for some kind of receptors corresponding to mapped 'cold spots' as units.

Seriatim mapping should be an ideal procedure for localizing 'cold spots,' because it provides statistical evidence of the reliability of reporting, which is not possible with single mapping. Isolated positive squares, surrounded by consistent zero squares, should be a feature of seriatim maps made with a 1 mm diam. stimulator—the traditional size for mapping 'cold spots.'

How does this theory fit the experimental facts? Figure 7 shows seriatim 1 mm maps from 10 highly consistent subjects—reliability coefficients ranging from .86 to .98. The left-hand maps give the actual seriatim scores. In spite of wide individual differences in general level, these maps show one common feature: *There are no discrete 'cold spots.'* There are consistent positive squares and consistent zero squares, but no isolated positive squares surrounded by zero squares, except in a few dubious instances. As a general rule, *the positive squares are grouped in clusters.*

It is idle to suggest that each square in a cluster represents a separate 'cold spot.' There is nothing sacred about these particular squares—the result of an arbitrary pattern. Stamping a slightly different pattern would give a new set of squares and require the assumption of a new set of 'cold spots' to fit. The number of 'cold spots' would then be indeterminate.

The Resolving Limits of Mapping.—Logically, the next step is to try smaller and smaller stimulators until resolution of the clusters into their component units is attained. Practically, this involves a number of difficulties. With a 1 mm diam. stimulator, we have reached about the limit of accurate

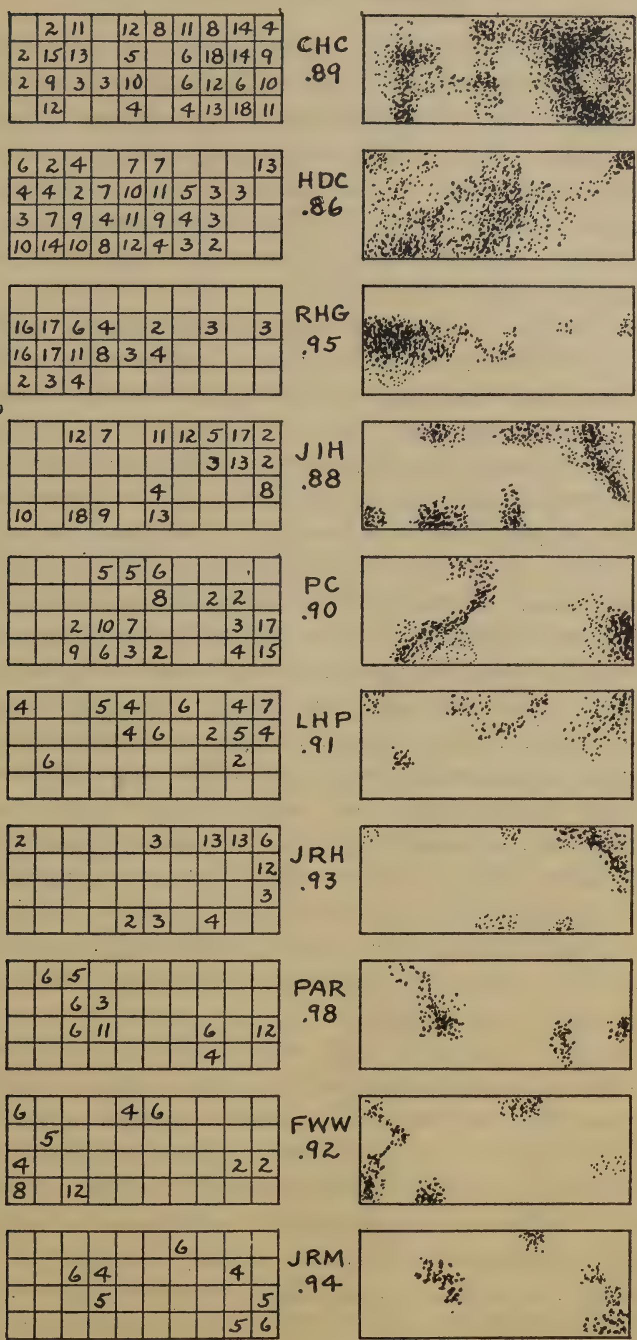


FIG. 7. Seriatim maps with 1 mm diam. stimulator. Each square in maps at left represents 1 mm². Maps at right illustrate the concentration theory.

manual placement. This might be overcome by the use of mechanical stimulation, if we could find some way of immobilizing the skin for a two-hour mapping period—no small task in itself. Again, the heat capacity of very small stimulators might set a limit on usable size. Even more important, the *spread* of stimulation (roughly proportional to the perimeter) might defeat our efforts to obtain a stimulator *functionally smaller* than a certain size.

Even if we could overcome these technical difficulties, however, there is nothing in the available evidence to make us expect that resolution can ever be attained. The maps in Fig. 7 are arranged in descending order of the number of positive squares. The upper two resemble those usually obtained with larger stimulator sizes, and the descending series roughly recapitulates the change in the appearance of the maps which occurs when we go from 3 × 3 to 2 × 2 to 2 × 1 to 1 mm diam. As we go down the series of maps, it is evident that the percentage of positive squares is becoming less and less—yet the remaining positive squares are still grouped into clusters. The lower maps are simply *disintegrating without being resolved into ultimate units.* Extrapolation of the process offers little hope that the end-point will be anything more than artifacts. Conceivably, we might reduce the stimulator size until only a single tiny point remains responsive out of each cluster. This, however, would not be a resolution of the cluster into component units, but merely a delimitation of the 'peak' of the cluster—no more indicative of the topography of the region than the last church spire which remains visible above flood waters.

Thermal-Conduction Inadequate.—Such 'peaks' could possess intrinsic significance only if each corresponded to a single receptor which is responsible for the variations of sensitivity found by mapping with somewhat larger stimulators. Thus, a *cluster* might represent a single receptor (or several of them)—the surrounding gradations being determined by thermal-conduction. Hardy and Oppel [8] have

[8] Hardy, J. D. and Oppel, T. W., Studies in temperature sensation. III. The sensitivity of the body to heat and the spatial summation of the end-organ responses, *J. clin. Investig.*, 1937, 16, 533–540.

suggested such a mechanism for warmth, thermal conduction to the receptor occurring through a network of blood vessels, making the receptor effective over much more than its actual area. This could account for two different part scores with a single end-organ as the intermediary. The critical test is the relation of whole to part scores. When a larger stimulator is applied to cover both parts simultaneously, in terms of the thermal-conduction hypothesis, the resulting intensity must be at least that of the maximum of the parts. A single receptor which is responsible for two part scores cannot conceivably be stimulated less strongly by a stimulator which covers both parts.

Yet this is exactly what does happen experimentally! The whole score, neglecting the minor size-effect, is the *square-root function* of the differing part scores, which is mathematically always *less* than the maximum. The establishment of the square-root function effectively shuts out thermal-conduction as an explanation of the variations in sensitivity found in seriatim maps. Hence, neither the traditional 'cold spots' nor thermal-conduction to isolated receptors can describe the distribution of sensitivity to cold.

The Concentration Theory.—How can we account for a distribution which is non-uniform, yet neither mappably punctiform, nor dependent upon variations in thermal-conduction? Let us assume that there are a great many minute receptors in a given skin area. These may vary individually in threshold and other characteristics—but fundamentally their *concentration* determines the sensitivity of the area. Where the receptors are densely packed, there is high sensitivity; where they are sparse, there is little or none. The situation is crudely represented by the shaded maps on the right-hand side of Fig. 7, deeper shading indicating a heavier concentration of receptors.

In terms of this theory, the intensity of cold experienced from a stated temperature is determined by two factors:

1. The *sensitivity* of the stimulated area—proportional to the square-root function of the concentration, or mathematically, the square root of the total number of receptors divided

by the area. With stimulators approximately 2 mm^2 and larger, this, and not size, is the primary determinant.

2. The *size* of the stimulated area—proportional mathematically to the total number of receptors stimulated. This is relatively unimportant with larger stimulators, but may exercise considerable influence below about 2 mm^2.

The theory, as above stated, is somewhat primitive. We may hope ultimately to reduce it to more basic neurological concepts; for example to the frequency of nerve impulses. Thus, the sensitivity factor might be translated into the *square root of the total frequency per unit area*, where 'total frequency' means the number of nerve impulses per second reaching the central nervous system from the area. The size factor would then become some fraction of the total frequency itself. We should like also to be able to identify the hypothetical minute receptors with specific anatomical entities; for example, with some of the 'free nerve endings,' usually assigned to pain by a process of exclusion. We might even hope eventually to discover some sort of differentiation which would constitute a distinguishing feature for those subserving cold sensitivity. All of these are goals for further research.

Broader Implications.—These findings with cold immediately raise the question: Is warm sensitivity punctiform in the traditional sense? Do 'warm spots' designate unit receptors any better than 'cold spots'? Seriatim maps of warm sensitivity—made with 2 and 3 mm stimulators—are similar in appearance to those for cold.[9] It seems equally improbable that the distribution can be adequately explained on the basis of 'warm spots'—especially as the number usually given is only one or two per cm^2. Thermal conduction from a single receptor cannot be used to explain the variations any more than with cold, unless it turns out that whole scores are equal to or greater than the maximum of part scores. Part-whole relations in seriatim warm-mapping will provide the critical evidence.

[9] Jenkins, W. L., Studies in thermal sensitivity. 10. The reliability of seriatim warm-mapping with untrained subjects, *J. exper. Psychol.*, 1939, **24**, 439–449.

Finally, some of the results suggest a possible extension beyond the field of temperature sensitivity. Four broad questions might be proposed:

1. Is the concentration theory a fundamental rule for all forms of cutaneous sensitivity?

2. Are all cutaneous receptors to be found among the 'free nerve endings,' with terminal differentiations for the various modalities?

3. Can the square-root function be extended into a general neurological law? That is, is the intensity of experience fundamentally proportional to the *square root of the total frequency of nerve impulses per unit area* in all sensory modalities—a modification of Adrian's frequency hypothesis? (In addition, there would be the minor size-effect—some fractional function of total frequency itself, a modification of Lucas' 'number of neurons active.')

4. Would the square-root function, if found translatable into nerve impulses, express a psychophysical relation of the kind Fechner was seeking? In place of measurements of external physical energy applied, we could employ the more direct stimulus units: frequency of nerve impulses. For the dubious direct measurement of sensation units, we could substitute seriatim scores of known reliability. The square-root function would then correspond to the Fechnerian concept of a logarithmic relation between stimulus and sensation.

These broad problems run far beyond the limits of the present data and are presented solely as a spur to further research.

Summary

Seriatim cold-mapping of the same cutaneous region alternately with two stimulators, the smaller being one-half or one-third the area of the larger, shows that scores indicating intensity of experience are dually determined: (1) There is a minor effect of size as such, relatively unimportant except with very small stimulators. (2) The major influence is the *square-root function*—whole scores being equal to the square root of the mean of the squares of differing part scores. Criti-

cal checks show that this square-root function holds with all magnitudes of part score differences and with all size comparisons tested.

Highly reliable seriatim maps with a 1 mm diam. stimulator reveal no isolated 'cold spots'; *i.e.*, positive squares surrounded by zero squares. The positive squares are always grouped in clusters, even in maps where very few of them remain. It is maintained that mapping with smaller stimulators cannot resolve these clusters into ultimate units, because the maps are simply disintegrating with no sign of resolution. The gradations within a cluster cannot be explained on the basis of varying thermal-conduction to a single receptor, because this would require whole scores equal to or greater than the maximum of part scores, whereas the relation follows the square-root function (necessarily less than the maximum). Neither the traditional 'cold spot' theory, nor the thermal-conduction modification provides an adequate explanation of the experimental facts.

In terms of a proposed *concentration theory*, sensitivity is determined primarily by the concentration of minute receptors. Neglecting a minor effect of size as such, the intensity of experience from stimulation with a stated temperature is proportional to the square root of the number of receptors per unit area, which might be translated neurologically as the square root of the total frequency of nerve impulses per unit area. The question is raised whether the concentration theory may be a fundamental rule for all forms of cutaneous sensitivity, and also whether the square-root function may express a general relation between intensity of experience and frequency of nerve impulses. These are problems for further research.

(Manuscript received May 2, 1939)

Reprinted from the JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
Vol. 25, No. 5, November, 1939

STUDIES IN THERMAL SENSITIVITY: 13. EFFECTS OF STIMULUS-TEMPERATURE IN SERIATIM COLD-MAPPING

BY WILLIAM LEROY JENKINS

Lehigh University

In the preceding study,[1] it was shown that the intensity of cold experienced from a given temperature depends primarily upon the sensitivity of the stimulated area and not upon the stimulus-size as such. The traditional 'cold spot' doctrine was found inadequate to explain the results, and a new *concentration theory* was proposed. In terms of this theory, sensitivity is a function of the concentration of relatively minute receptors; specifically, it is proportional to the square root of the number of receptors per unit area.

The question naturally arises: Can this theory be extended to cover the effects of stimulus-temperature as well as those of stimulus-size? An answer to this question can be approached by obtaining comparative seriatim scores with a standard stimulator size at different temperatures.

PROCEDURE

The seriatim mapping was carried out in substantially the same manner as in the previous study.[2] One member of each group acted as experimenter; the other two alternated as subject and recorder. At each experimental period a checkerboard of 3 × 3 mm squares was mapped 12 times, alternately with a standard and a comparison temperature. The standard of 17° C. was compared with 8°, 11°, 20°, 23° and 26° C., and the standard of 23° C. with 20°, 26°, 27½° and 29° C.—all temperatures being measured at the inlet. The mapping followed an approximately random sequence, as previously described.

Subjects reported 0 for neutral, 1 for weak, 2 for medium and 3 for strong cold. The *score* for each square was determined by adding up the 6 numbers reported. The empirical justification for treating these ordinal numbers as if they were cardinal has been discussed in the previous article.[3] Correlations between scores at standard and comparison temperatures were computed by the product-moment formula.

[1] Jenkins, W. L., Studies in thermal sensitivity: 12. Part-whole relations in seriatim cold-mapping, *J. exper. Psychol.*, 1939, 25, 373–384.

[2] *Op. cit.*

[3] *Op. cit.*

Results

The upper part of Table 1 shows the correlations obtained for all subjects. At temperatures of 23° and lower, approximately half were .80 or better, which accords with the results of the preceding study.[4] At 26°, 27½° and 29°, however, there was a progressive reduction until at 29° C. no subject attained a correlation as high as .80.[5] To minimize the

TABLE 1

Correlations

All Subjects

Temp. ° C.	.00 .59	.60 .64	.65 .69	.70 .74	.75 .79	.80 .84	.85 .89	.90 .94	.95 1.00	Total Cases	Med. Corr.
17/8	1	2	–	5	4	3	3	2	–	20	.77
17/11	4	1	2	1	2	4	3	3	–	20	.80
*17/17	4	–	1	5	8	14	5	2	1	40	.81
17/20	4	–	1	5	3	2	2	1	–	18	.74
17/23	2	1	3	3	1	6	5	1	–	22	.81
17/26	4	2	3	2	–	1	4	–	–	16	.68
23/20	3	–	1	2	1	2	4	1	–	14	.80
23/26	1	1	1	4	4	2	–	1	–	14	.75
23/27½	10	1	2	3	2	1	–	–	–	19	.60
23/29	14	6	6	2	2	–	–	–	–	30	.60

Selected Subjects

Temp. ° C.	.00 .59	.60 .64	.65 .69	.70 .74	.75 .79	.80 .84	.85 .89	.90 .94	.95 1.00	Total Cases	Med. Corr.
17/8	–	1	–	2	3	3	2	2	–	13	.81
17/11	–	–	1	–	1	3	3	3	–	11	.86
*17/17	2	–	–	2	4	8	4	1	1	22	.82
17/20	–	–	–	–	1	1	2	1	–	5	.87
17/23	1	–	–	–	1	4	4	1	–	11	.84
17/26	2	1	2	1	–	1	3	–	–	10	.70
23/20	–	–	–	1	–	1	4	1	–	7	.86
23/26	–	–	1	3	3	2	–	1	–	10	.77
23/27½	3	–	2	3	2	1	–	–	–	11	.71
23/29	5	6	2	2	1	–	–	–	–	16	.63

* Preliminary check on reliability, performed on the first day, without previous training.

[4] *Op. cit.*

[5] Most of the scores at 29° C. were so low as to fall below the level of accurate discrimination, even for subjects who showed excellent consistency at lower temperatures. It is significant, however, that two subjects with relatively high average scores gave odd-even correlations of .85 and .92 in the 29° reports, although the correlations of 29° with 23° scores were below .80. This indicates that the behavior of the scores themselves may become irregular as the limen is approached.

effects of random reporting, a group of the better subjects was selected on the basis of their general consistency at lower temperatures, and only their results used in the later computations. The correlations for these selected subjects are shown in the lower part of Table 1.

As a preliminary step, *average scores* were computed for each stimulus-temperature. In Fig. 1, the large chart shows the mean scores for all subjects combined; the small charts give similar results for the individual subjects. Each dot in the latter represents the mean score from a map of 50 squares. Only two temperatures could be used at each weekly meeting and a new map had to be stamped on the skin each time. In spite of the fact that the data were thus collected over a period of five weeks, there is a surprising amount of consistency shown in the average scores on the majority of the charts.

A comparison of the large chart with the small ones shows that the form of the combined curve is a compromise obtained by adding a number of variant forms. Yet the general shape of this curve is reproduced in the majority of the charts (*i.e.*, A, C, D, F, G, I, J, L, N, S, T, W). The variants range from a horizontal line with no clear evidence of a break (K), through descending curves (B, M, Q, R) to sloping lines (E, H). A few cases (O, P, U, V) are too irregular for interpretation. Considering only the most typical form, it appears that in these subjects temperature changes below 23° C. have little or no effect on average scores.

If we now plot *individual* scores at standard temperatures against those at comparison temperatures, the graphs shown in Fig. 2 are obtained. Here the data for all 23 selected subjects have been combined in order to secure evidence on the general effects of temperature at various score levels. The ordinates represent scores at standard temperatures, grouped into 7 divisions for clarity. The abscissæ represent scores at comparison temperatures. The circles show median comparison scores corresponding to the seven standard score levels. Roughly, these charts show what might be expected from the average score curves. Below 23° C., the circles are close to

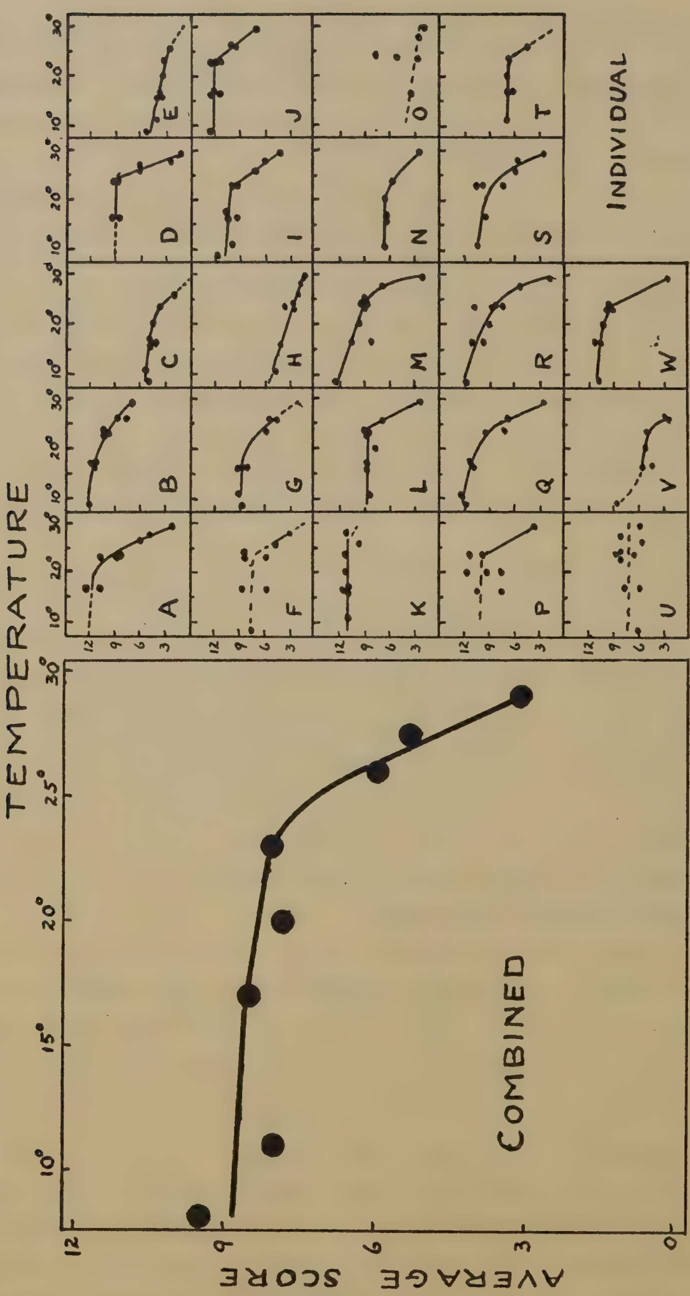


FIG. 1. Relation of average scores to stimulus-temperatures.

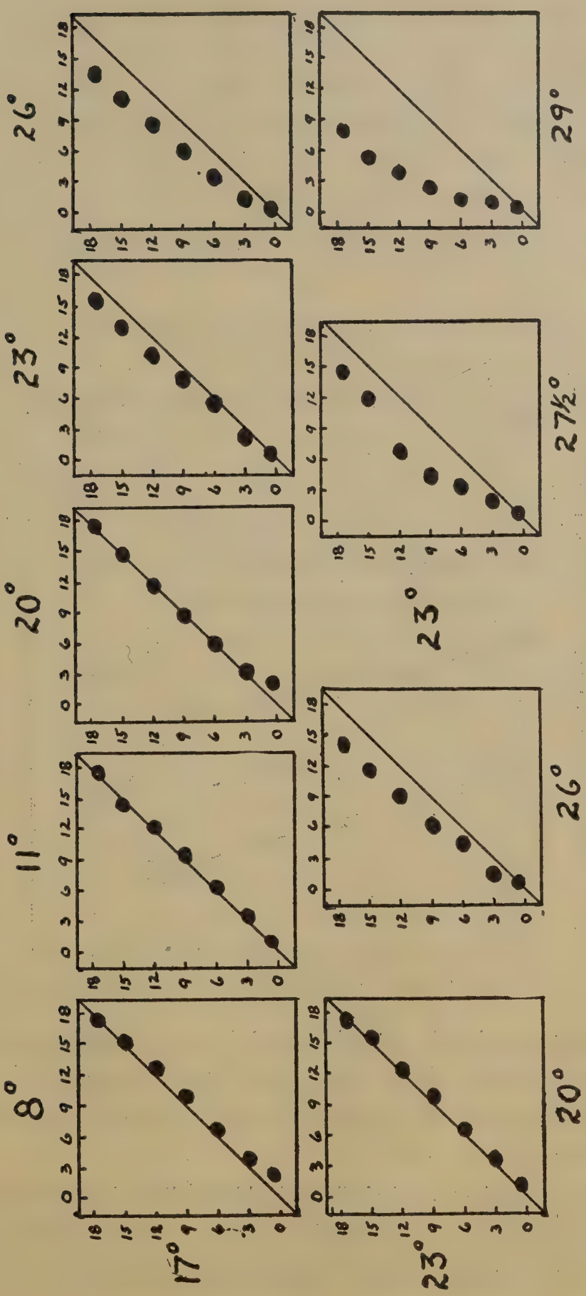


FIG. 2. Median comparison scores for 7 standard score levels.

the diagonal. At 26°, 27½° and 29° C., there is progressive divergence.[6]

A closer inspection, however, shows that the divergence from the diagonal is *relatively* greater at the lower score-levels. To investigate this relation more exactly, the *percentages* of change were computed with the 17° scores as 100, making the necessary adjustments where the actual standard was 23° C. These percentages are shown graphically in Fig. 3. They

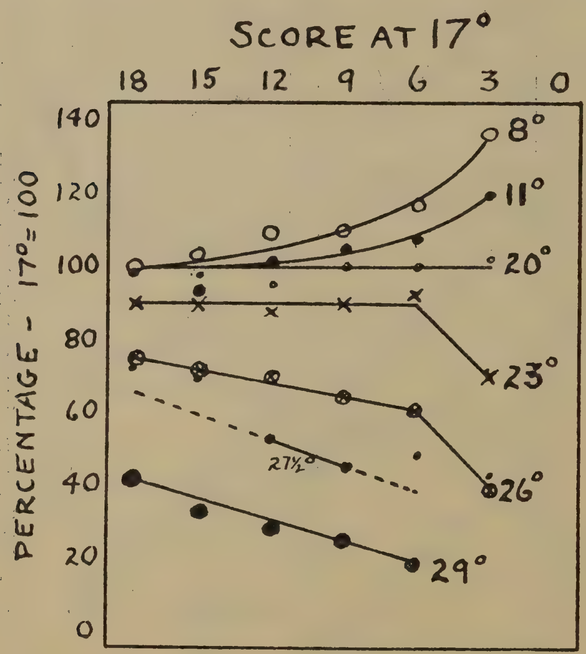


FIG. 3. Percentage changes with temperature for 7 standard score levels.

follow a systematic distribution which can be expressed in the following generalization: *Both above and below 17° C., the percentage change with temperature is progressively greater at the lower score levels.*

[6] Likewise, similar charts for the individual subjects follow what might be expected from the individual charts of average scores. If the average score increases below 23° C., the medians show a corresponding divergence to the right. If the average score is relatively unchanged at lower temperatures, the medians are close to the diagonal.

These curves are based on the combined data of 23 subjects, neglecting possible individual differences. It is safe to say, however, that the principle applies at least to the majority of the subjects. Detailed examination of the individual scattergraphs fails to disclose a single clear-cut exception.

Discussion

On the basis of the foregoing results, three new hypotheses can be proposed:

1. *Individual Cold Receptors Are All-or-none in their Response.*—The evidence for this is Fig. 1. Many of the individual charts show approximately a horizontal line between 8° and 23°, meaning that the average score is unchanged through this range. It follows that the individual receptors in these subjects must obey the same law; *i.e.*, they must be unaffected by supra-liminal temperature changes.

It is unimportant whether the receptor action rises instantly, or merely very rapidly, to the full peak. Likewise, it is evident that the peak might vary from time to time as the result of adaptation or other changes. The major point is that once this peak is reached, the activity is unaffected by further temperature change, and is all-or-none in this sense.

2. *The Concentration Theory can be Extended to Cover the Effects of Temperature.*—When a lower temperature causes a more intense experience it does so merely by adding to the concentration of active receptors. This follows automatically if we assume that cold receptors are all-or-none in all subjects. The various curves in Fig. 1 then become dependent upon the distribution of temperature limens.[7]

3. *The Distribution of Temperature Limens among Receptors Is not Random.*—If the distribution were purely random, a change of temperature should cause the addition or subtraction of the same percentage of active receptors at all score-levels. The lines of Fig. 3 would then be parallel to the horizontal axis. However, the lines fan out to the right,

[7] It is possible to assume that cold receptors are all-or-none in some subjects and not in others. In the absence of direct evidence, parsimony dictates the assumption of the simpler hypothesis: the same kind in all individuals.

indicating that temperature changes cause progressively greater percentage changes in score at the lower score-levels. In terms of the concentration theory, this means that the distribution of temperature limens is different in low-score and high-score regions. In low-score regions there are relatively more unresponsive receptors (*i.e.*, those that respond only to low temperatures). In high-score regions, there are relatively more highly-responsive receptors. This is significant, because it suggests that the distribution of *strength* (*i.e.*, frequency of nerve impulses transmitted) may not be random among receptors, but may follow an analogous plan. That is, high-score regions may contain relatively more strong and highly-responsive receptors; low-score regions relatively more weak and unresponsive receptors. *Concentration* and *strength* would then work hand-in-hand.

If the distribution of strength among receptors is purely random, there is only one variable to determine sensitivity: *number*. Because of the square-root function, the number of receptors would have to be very large. For example, a sensitivity range from 1 to 10 would require receptor concentrations from 1 to 100. (Even with fairly small stimulators, it seems probable that we can establish close to 10 score-levels in a single map that are reliably different statistically.) On the other hand, with *two* variables—number and strength—all of the gradations could be accounted for in terms of a much smaller number of receptors. Sensitivity would be proportional, not to the square root of the number of receptors, but to the square root of some more fundamental neurological units, probably nerve impulse frequency. The greater flexibility of this revised hypothesis recommends it for provisional adoption.

Summary

An investigation of comparative seriatim scores at different stimulus-temperatures leads to the formulation of three hypotheses:

1. *Cold Receptors Are All-or-none in their Response.*—Average scores in many subjects show no change between 23° and 8°; consequently the individual receptors (in these

subjects, at least) must be unaffected by supra-liminal changes of stimulus-temperature.

2. *The Concentration Theory can be Extended to Cover the Effects of Stimulus-Temperature.*—When a lower temperature causes a more intense experience, it does so merely by adding to the concentration of active receptors. This follows automatically unless it is assumed that cold receptors are all-or-none in some subjects and obey a different principle in others.

3. *The Distribution of Limens among Receptors Is not Random.*—Low scores show greater percentage changes with temperature than high scores. Therefore, there must be relatively more poorly-responsive receptors in low-score regions; relatively more highly-responsive receptors in high-score regions. This suggests that the distribution of *strength* among receptors may also not be random, but that there may be relatively more strong receptors in high-score regions; relatively more weak receptors in low-score regions. Concentration and strength would then work hand-in-hand. In this more flexible interpretation of the concentration theory, sensitivity would be determined by the square root of the concentration of some fundamental neurological units (such as nerve impulse frequency), and the number of receptors required would be greatly reduced.

(Manuscript received May 13, 1939)

DATE DUE
